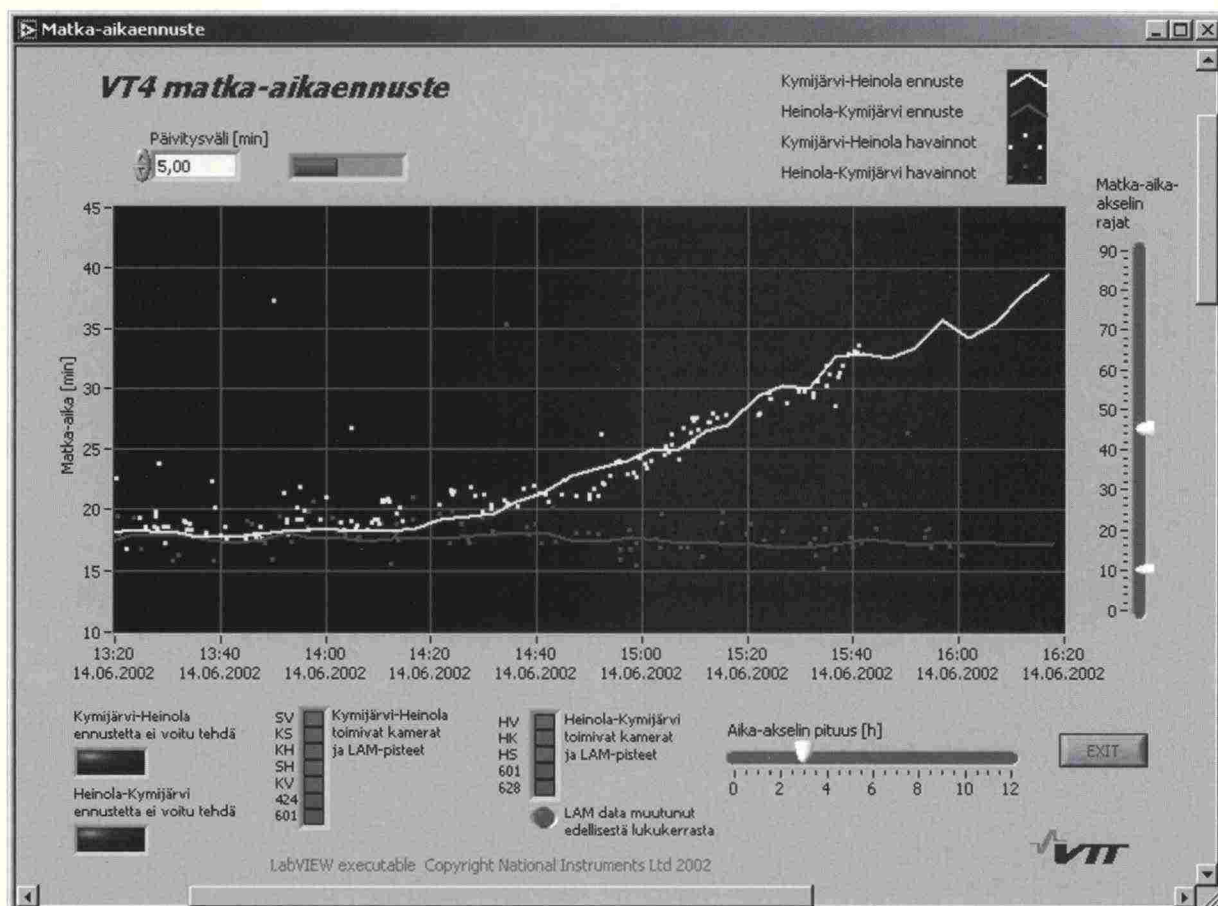


Matka-ajan lyhyen aikavälin ennustemalli

Kokeiluversio Lahti-Heinola-välille

Tiehallinnon selvityksiä 22/2002



Satu Innamaa

Matka-ajan lyhyen aikavälin ennustemalli

Kokeiluversio Lahti-Heinola-välille

Tiehallinnon selvityksiä 22/2002

ISSN 1457-9871
ISBN 951-726-897-1
TIEH 3200754

Edita Prima Oy
Helsinki 2002

Julkaisua myy/saatavana:
Tiehallinto, julkaisumyynti
faksi 0204 22 2652
s-posti julkaisumyynti@tiehallinto.fi
www.tiehallinto.fi/julk2.htm



TIEHALLINTO
Liikenteen palvelut
Opastinsilta 12 A
PL 33
00521 HELSINKI
Puhelinvaihte 0204 22 11

Satu Innamaa: Matka-ajan lyhyen aikavälin ennustemalli. Kokeiluversio Lahti-Heinola-välille. Helsinki 2002. Tiehallinto, Liikenteen palvelut. Tiehallinnon selvityksiä 22/2002. 32 s. + liitt. 16 s. ISSN 1457-9871, ISBN 951-726-897-1, TIEH 3200754.

Asiasanat: liikennetiedotus, ennusteet, matka-aika
Aiheluokka: 11, 20

TIIVISTELMÄ

Matka-aikatiedotuksella voidaan vaikuttaa tienkäyttäjien matka-aikaan kohdistuviin odotuksiin ja siten tavoitematka-aikoihin. Täten pyritään parantamaan koettua sujuvuutta.

Aikaisemmassa tutkimuksessa kehitettiin valtatielle 4 Lahden ja Heinolan välille matka-ajan lyhyen aikavälin ennustemalleja. Ennustemalleista saatujen hyvien kokemusten perusteella malleista päätettiin tehdä kokeiluversiot. Aikaisemmat ennustemallit tehtiin ottamatta huomioon todellisten järjestelmien aiheuttamia ylimääräisiä viiveitä ynnä muita rajoituksia ajantasaisessa käytössä. Nyt tehdyissä kokeilumalleissa nämä todellisen elämän rajoitukset otettiin huomioon.

Ennustemalleina käytettiin monikerrosperspeptronineuroverkkoja. Maastosta tulevia matka-aikahavaintoja ei suodatettu havaintojen vähäisyyden takia. Syötemuuttajaehdokkaina olivat eri osalinkkien matka-aikamediaanit sekä LAM-pisteissä mitatut liikennemäärät ja keskinopeudet. Mallien syötesuureet valittiin eri ehdokkaiden joukosta sillä perusteella, kuinka hyvin ne korreloivat ennustettavan matka-ajan kanssa ruuhkassa.

Etenkin matka-aika-aineistossa on paljon lyhyitä katkoksia. Mallin pitäisi toimia mahdollisimman joustavasti katkoksista huolimatta. Lyhyisiin katkoksiin varauduttiin kehittämällä päivityssääntö, jonka avulla kunkin syöteen arvoa päivitetään. Pidempiin katkoksiin varauduttiin opettamalla neuroverkkoja ilman joidenkin ilmaisimien havaintoja.

Päätettiin, ettei ennustetta kuitenkaan tehdä, jos hyvän ennusteen tekemiseen ei ole riittävästi edellytyksiä. Jos liian moni tai ennusteen kannalta olennainen ilmaisin on poissa toiminnasta, malli ei anna ennustetta.

Malleja verrattiin parhaisiin aikaisempiin malleihin. Tulokset olivat yhtenevät nyt tehtyjen mallien kanssa, joten niihin voi olla tyytyväinen. Malleja voidaan lisäksi pitää varsin hyvin varautuneina yksittäisten ilmaisimien rikkoutumisiin.

Malleja voitaisiin parantaa jonkin verran mittaamalla linkille saapuvaa liikennemäärää ja kasvattamalla matka-aikahavaintojen otoskokoa.

Satu Innamaa: Matka-ajan lyhyen aikavälin ennustemalli. Kokeiluversio Lahti-Heinola-välille. [Short-term prediction model for travel time. Pilot version for Lahti-Heinola section.] Helsinki 2002. Finnish National Road Administration. Finnra Reports 22/2002. 32 p. + app. 16 p. ISSN 1457-9871, ISBN 951-726-897-1, TIEH 3200754

Keywords: road user information, predictions, travel time

SUMMARY

The road users' travel time expectations and thereby their objective travel time can be influenced by travel time information. This way, the experienced fluency is tried to be improved.

In a previous study, short-term prediction models of travel time were developed for the section between Lahti and Heinola on Finnish main road 4. Due to the good experience obtained from the prediction models, a pilot version was to be made. The previous models were made without taking into account the extra delays and other limitations caused by various systems when used in real time. In the pilot models made in this study, the limitations of real life were taken into account.

Multilayer perceptron neural networks were used as prediction models. The travel time observations coming from the field were not filtered because of the small sample size. Input variable candidates were median travel times for different sub links, and the flow and the mean speed measured in the automatic traffic measurement points (so called LAM-points). The input variables were selected from the candidates according to their correlation with the travel time to be predicted in congested conditions.

Especially the travel time data had several short breaks. The model should work as fluently as possible despite of them. The models were prepared for short breaks by developing an updating rule according to which the value of an input variable is updated. The models were prepared for longer breaks by training neural networks without observations from certain detectors.

However, it was decided that a forecast is not made, if the requirements for making a good forecast are not met. If too many detectors or one essential to the forecast do not function, the model does not make a forecast.

The models were compared with the best models from the previous study. The previous results were congruent with the models developed in this study, so the current models were satisfactory. In addition, the models can be considered to be rather well prepared for the breakdown of individual detectors.

The models could be improved by monitoring the flow that is coming to the link and by increasing the sample size of the travel time observations.

ESIPUHE

Tiehallinto sekä liikenne- ja viestintäministeriö ovat tilanneet tutkimuksen automaattisista liikenteenohjaus- ja liikenteen tiedotusjärjestelmistä osana ministeriön FITS – Liikennetelematiikan rakenteiden ja palvelujen tutkimus- ja kehittämisohjelmaa 2001–2004. Tehty tutkimus matka-ajan lyhyen aikavälin ennustemalleista on valmistunut osana kyseistä toimeksiantoa.

Tutkimuksesta vastasi tutkija Satu Innamaa VTT Rakennus- ja yhdyskuntatekniikasta. Mallien lopullisen ohjelmoinnin ja ennusteohjelman käyttöliittymän teki tutkimusinsinööri Mikko Kallio VTT:sta.

Tutkimuksen etenemistä valvoi asiantuntijaryhmä, johon kuuluivat liikenne-neuvos Matti Roinne liikenne- ja viestintäministeriöstä, kehitysinsinöörit Sami Luoma ja Jyri Vilhunen Tiehallinnon liikenteen palvelut –yksiköstä, liikenne-keskusvastaava Mauri Pyykönen ja liikenteen palvelupäällikkö Jyri Mustonen Uudenmaan tiepiiristä, professori Matti Pursula ja opettava tutkija lisäksi Kossonen Teknillisen korkeakoulun liikennelaboratoriosta sekä tutkimusprofessori Risto Kulmala VTT:sta.

Neuraalilaskennassa on hyödynnetty CSC - Tieteellinen Laskenta Oy:n myöntämiä resursseja.

Helsinki, kesäkuussa 2002

Tiehallinto
Liikenteen palvelut

Sisältö

1	TAUSTAA	9
2	TAVOITE	10
3	TUTKIMUSKOHDE	11
4	MLP-NEUROVERKKOMALLIT	13
5	AINEISTO	15
5.1	Reunaehdot	15
5.2	Aineisto	16
6	ENNUSTEMALLIT	17
6.1	Syötesuureiden valinta	17
6.2	Mallien rakentaminen	18
7	ENNUSTEMALLIEN HYVYYS	20
8	MALLIEN OHJELMOIMINEN	24
9	JOHTOPÄÄTÖKSIÄ JA SUOSITUKSIA	29
	LÄHDELUETTELO	31
	LIITTEET	32

1 TAUSTAA

Yksilötasolla tieosuuden tai muun liikennejärjestelmän osan liikenteen sujuvuutta arvioidaan sen perusteella, kuinka häiriöttömästi ja odotusten mukaisesti tienkäyttäjä kykenee kulkemaan sen läpi. Tärkein yksittäinen objektiivinen liikenteen sujuvuuden mittari on matka-aika. Odotusten mukaisuudella kuvataan matkan ennustettavuutta. Liikkujien kannalta olennaista näyttää olevan kokonaismatka-ajan ohella se, kuinka paljon joudutaan poikkeamaan oletetusta tai tavoitematka-ajasta. (Luoma 1998.)

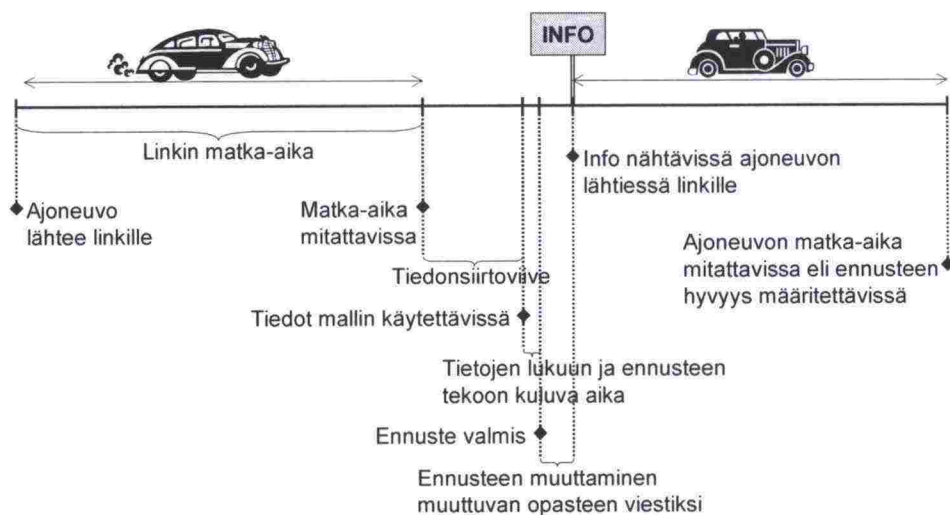
Kertomalla ajantasaisesti liikennetilanteesta yritetään tienkäyttäjää auttaa valitsemaan liikkumisen kannalta parhaat ajankohdat ja reitit. Näin tieverkon välityskykyä käytetään tehokkaasti hyväksi. Ennen matkaa saatu tieto vaikuttaa lähinnä reitin valintaan, kun taas matkan aikana saatu tieto vaikuttaa ajotavan muuttamiseen. (Kiljunen & Summala 1996.)

Ennustettu matka-aika on liikenneoperaattorin näkökulmasta katsottuna erittäin tärkeä, sillä juuri se vaikuttaa päätöksentekoon tienkäyttäjille tarjottavasta tiedosta vaihtoehtoisista, paremmista ajoreiteistä. Tiedotuksella pyritään vaikuttamaan tienkäyttäjien matka-aikaodotuksiin ja siten tavoitematka-aikoihin. Täten pyritään parantamaan koettua sujuvuutta. (Luoma 1998.)

2 TAVOITE

Aikaisemmassa tutkimuksessa (Innamaa ym. 2002) kehitettiin valtatielle 4 Lahden ja Heinolan välille matka-ajan lyhyen aikavälin ennustemalleja. Ennustemalleista saatujen hyvien kokemusten perusteella malleista päätettiin tehdä kokeiluversiot, jotka tuottavat ennusteita tienkäyttäjille edelleen annettaviksi. Tämä selvitys raportoi ennustemallista tehdyn kokeiluversion toteutuksen.

Aikaisemmat ennustemallit tehtiin ottamatta huomioon todellisten järjestelmien aiheuttamia ylimääräisiä tiedonsiirto- ynnä muita viiveitä (Kuva 1) olettamalla, että kaikki tieto on heti mallin käytettävissä. Samoin niissä oletettiin, että raaka-aineistosta on mahdollista suodattaa poikkeavat havainnot pois, ennen kuin se muokataan erilaisiksi matka-aikajakaumaa kuvaaviksi tunnusluvuiksi. Todellisuudessa näin ei kuitenkaan välttämättä ole. Nyt tehtävissä kokeilumalleissa todellisen elämän rajoitukset on otettava huomioon, eivätkä tulokset välttämättä ole kaikilta osin yhtä hyviä kuin ihanteellisiin olosuhteisiin tehdyillä malleilla. Mallien hyvyttä verrataankin luvussa 7.



Kuva 1. Aikajana, joka kuvaa erilaisia viiveitä siitä hetkestä, kun se ajoneuvo, jonka matka-aikaan ennuste perustuu, lähtee linkille, aina siihen hetkeen, kun matka-aikatiedon nähnyt kuljettaja pääsee linkin loppuun ja voi arvioida matka-aikatiedon oikeellisuuden.

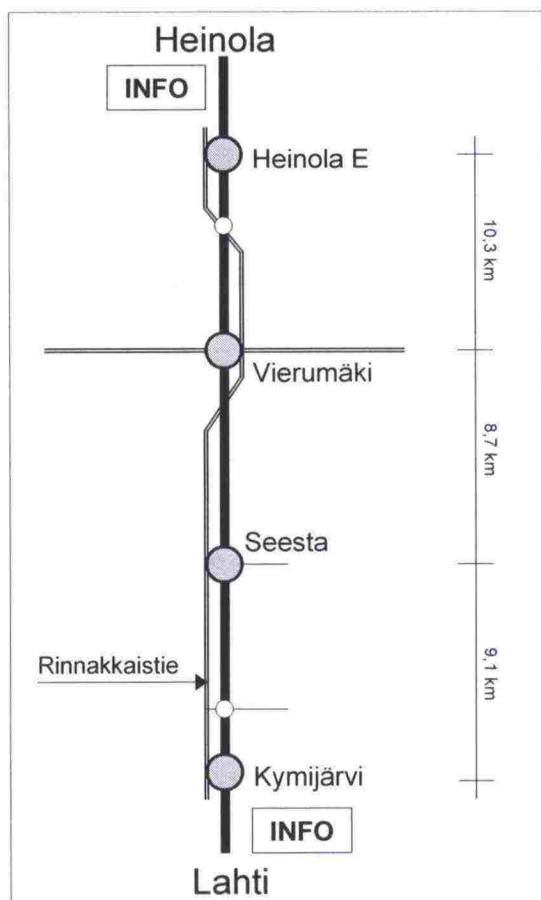
3 TUTKIMUSKOHDE

Tutkimuskohteena oli valtatie 4 Lahden ja Heinolan välillä. Tieosalla kesä-ajan keskimääräinen vuorokausiliikennemäärä on noin 15 000 ajon./vrk (Tielaitos 1999).

Koejakso on varustettu matkanopeuden seurantajärjestelmällä, jonka avulla matka-ajat mitataan automaattisesti rekisteritunnusten perusteella. Tieosalla on muuttuvat opastetaulut, jotka kertovat odotettavissa olevan matka-ajan. Opaste sijaitsee molemmilla suunnilla ennen ensimmäistä matka-ajan mittauskameraa, noin 36 kilometriä ennen kohteena olevaa kaupunkia. Eteläpäässä (Lahti) opaste sijaitsee tarkemmin sanottuna noin 4 kilometriä ennen Joutjärven eritasoliittymää ja pohjoispäässä (Heinola) heti Heinolan tähden vesistö sillan jälkeen.

Matkanopeuden seurantajärjestelmän toiminta perustuu ajoneuvojen tunnistamiseen liikennevirrasta ilman ajoneuvoihin asennettavia erillisiä tunnistiteita tai lähettämiä. Ajoneuvo pyritään tunnistamaan yksilöllisesti rekisterikilven perusteella aina, kun se ohittaa järjestelmän mittauspisteen. Pisteen ohi kulkeneiden ajoneuvojen rekisteritunnukset ja ohitusajankohdat tallennetaan, ja peräkkäisten pisteiden tietoja verrataan keskenään. Yhdistämällä saman ajoneuvon tiedot kahdessa pisteessä saadaan laskettua ajoneuvon matkaan käyttämä aika eli ajoneuvon matka-aika. Järjestelmä perustuu Golden River Traffic Ltd:n toimittamiin laitteistoihin. (Eloranta 1999.)

Matka-ajan seurantajärjestelmän kamerapisteen sijaitsevat Kymijärvellä, Seestassa, Vierumäellä ja Heinola E:ssä (Kuva 2). Tieosalla on lisäksi yksi liikenteen automaattinen mittauspiste (LAM-piste), joka sijaitsee Vierumäellä. Vierumäen (LAM 601) lisäksi Lahden eteläpuolella (LAM 424) ja Heinolan pohjoispuolella (LAM 628) on LAM-pisteet.



Kuva 2. Tutkimusalue valtatiellä 4. Tutkimusalueella on kamerapisteet Heinolassa, Vierumäellä, Seestassa ja Kymijärvellä sekä LAM-piste Vierumäellä. Muuttuvat opastetaulut kertovat matka-ajasta tieosan molemmissa päissä. Kuvaan on merkitty kamerapisteiden väliset etäisyydet.

Kymijärven ja Heinolan kamerapisteet sijaitsevat moottoritiellä, vajaa kilometri ennen ohituskaistatien alkua tai ohituskaistatien loppumisen jälkeen. Näissä pisteissä kaistoja on kaksi molempiin suuntiin. Seesta ja Vierumäki sijaitsevat ohituskaistatiellä, jossa on peruskaistojen lisäksi yksi ohituskaista. Tämä ohituskaista on molempien mittauspisteiden kohdalla etelään eli Lahden suuntaan. Matka-aikaseuranta kattaa yhden kaistan (peruskaista) suuntaa kohti kussakin mittauspisteessä. Koska Seestassa ja Vierumäellä ei ole ohituskaistaa Heinolan suunnalla, kaikkia suunnan ajoneuvoja seurataan. Muuten seuranta kohdistuu ainoastaan peruskaistaa ajaviin ajoneuvoihin.

Ennustemallien kokeiluversiot tehtiin Kymijärvi–Heinola- ja Heinola–Kymijärvi-linkeille. Mallien antamia estimaatteja voidaan käyttää muuttuvilla opastetauluilla näytettävien matka-aikaestimaattien pohjana.

4 MLP-NEUROVERKKOMALLIT

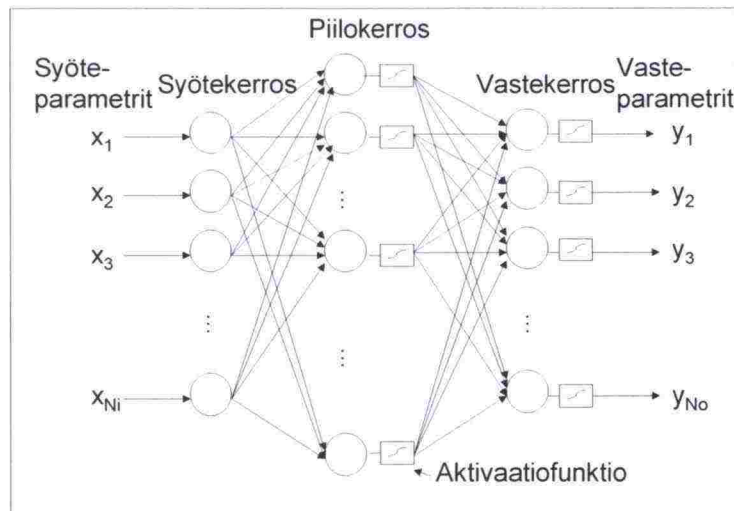
Ennustemalleina käytettiin monikerrospersptronineuroverkkoja (MLP-verkot, Kuva 3) niillä aikaisemmissa tutkimuksissa (Innamaa ym. 2002, Innamaa & Pursula 2000, Lee ym. 1998, Smith & Demetsky 1994 ja 1997) saatujen hyvien tulosten perusteella. Syötekerroksen neuronien määrä oli yhtä suuri kuin syöteparametrien määrä, ja vastekerroksen neuronien määrä vastaavasti yhtä suuri kuin vasteparametrien määrä. Piilokerrosten määräksi valittiin yksi. Piiloneuronien määrä määriteltiin siten, että opetusjoukon koko (T) oli vähintään viisi kertaa estimoitavien parametrien (paino- ja harhaker-toimet, W) lukumäärä. N_i on syöteneuronien lukumäärä, N_h piiloneuronien lukumäärä ja N_o vasteneuronien lukumäärä.

$$T_{\min} = 5 \cdot W$$

$$W = (N_i + 1)N_h + (N_h + 1)N_o$$

$$\Rightarrow N_{h,\max} = \frac{0,20 \cdot T - N_o}{N_i + N_o + 1}$$

Jos näin saatu piiloneuronien maksimimäärä oli yli 20, neuronien määräksi valittiin 20, jottei opetusprosessi käynyt liian raskaaksi.



Kuva 3. Monikerrospersptronineuroverkko.

Mallien opettamiseen käytettiin Fletcher-Reevesin päivitystä, joka kuuluu konjugaatti-gradienttimenetelmiin. Perusgradienttimenetelmissä neuroverkon painokertoimia päivitetään siihen suuntaan, jossa hyvyysfunktio (esimerkiksi virheen neliösumma) pienenee nopeimmin. Tämä ei kuitenkaan välttämättä takaa nopeinta konvergenssia. Konjugaatti-gradienttimenetelmässä suunnaksi valitaan konjugaatin suunta, jolloin päästään nopeampaan konvergenssiin kuin suurimman gradientin suunnassa. (Demuth & Beale 1998.)

Jos neuroverkko oppii opetusaineiston liian hyvin, se muistaa sen ulkoa eikä kykene yleistämään. Eräs tapa ehkäistä ulkoa oppimista on varmistaa, että opetusjoukko on riittävän suuri (piiloneuronien maksimimäärä perustuu opetusjoukon kokoon). Toinen tapa välttää ylioppiminen on asettaa opetusprosessille lopetuskriteerejä. Nyt tehdyssä tutkimuksessa näitä kriteerejä asetettiin useita. Opettaminen lopetettiin, kun saavutettiin opetuskierrosten maksimimäärä, kun gradientin arvo tai keskimääräinen neliövirhe laskivat riittävän pieniksi tai kun kalibrintiaineiston keskimääräinen neliövirhe lakka- si pienenemästä. Käytännössä opetus lopetettiin aina viimeisen lopetuskri- teerin perusteella.

Viimeksi mainittua kalibrintiaineistokriteeriä varten alkuperäinen opetusai- neisto jaettiin kolmeen osaan: opetus-, kalibrinti- ja validointiaineistoon. Opetusaineisto koostui niistä havainnoista, joiden avulla neuroverkolle yritet- tiin opettaa mallinnettava ilmiö. Kalibrintiaineiston perusteella yritettiin ha- vaita, milloin opetus oli edennyt niin pitkälle, ettei neuroverkko enää oppinut yleisiä mallinnettavaan ilmiöön liittyviä ominaisuuksia, vaan alkoi opetella käytetyn opetusjoukon erityispiirteitä. Tämä ilmeni tilanteena, jolloin opetus- joukon virhe pieneni, mutta kalibrintijoukon virhe alkoi kasvaa. Validointi- joukko oli kolmas aineisto eikä sitä käytetty opetusprosessissa millään ta- valla. Sen tarkoituksena oli testata, kuinka hyvin opetuksessa onnistuttiin eli kuinka hyvin neuroverkko toimi.

5 AINEISTO

5.1 Reunaehdot

Matka-aikatiedot saadaan maastosta viiden minuutin välein liikennekeskukseen. Tiedot tulevat ASCII-muodossa ADSL-yhteyden välityksellä. Kuvassa 4 on esimerkki ajoneuvokohtaisista matka-aikatiedoista. (Kokkinen 2002.)

tunnus	linkki	rekisteri	matka-aika	lähtöaika	tuloaika	lähtöpiste	tulopiste	lähtötunnus	tulotunnus
4488508	0	FI:XXXXXX	337	2002-01-14 02:40:47	2002-01-14 02:46:24	0	2	21805609	21805623
4488509	0	FI:XXXXXX	394	2002-01-14 02:41:33	2002-01-14 02:48:07	0	2	21805610	21805624
4488522	0	FI:XXXXXX	355	2002-01-14 04:11:53	2002-01-14 04:17:48	0	2	21805825	21805846
4488526	0	FI:XXXXXX	460	2002-01-14 04:22:04	2002-01-14 04:29:44	0	2	21805862	21805891
4488528	0	FI:XXXXXX	343	2002-01-14 04:27:36	2002-01-14 04:33:19	0	2	21805884	21805904
4488545	0	FI:XXXXXX	369	2002-01-14 04:55:21	2002-01-14 05:01:30	0	2	21806055	21806062
4488550	0	FI:XXXXXX	365	2002-01-14 05:02:15	2002-01-14 05:08:20	0	2	21806092	21806123
4488551	0	FI:XXXXXX	389	2002-01-14 05:03:51	2002-01-14 05:10:20	0	2	21806097	21806128
4488554	0	FI:XXXXXX	406	2002-01-14 05:08:46	2002-01-14 05:15:32	0	2	21806110	21806176

Kuva 4. Esimerkki liikennekeskukseen saapuvista matka-aikatiedoista (Kokkinen 2002).

Liikenneolosuhdetietojen keruuohjelma (tiesää-windows'in NT-keruu) tuottaa viiden minuutin välein viimeisimmälle viidelle minuutille laskettuja pistekoh-
taisia liikennetietoja (LAM-pisteittäin). Kuvassa 5 on esimerkki järjestelmän
tuottamasta tiedosta. Tiedot saapuvat järjestelmään 10–15 minuutin viipeel-
lä. (Vilhunen 2002.)

piste	pvä	h	min	q1	v1	q2	v2
101	23	12	15	139	87	131	86
102	23	12	10	49	96	43	92
103	23	12	10	12	108	83	101
107	23	12	15	96	88	91	86
109	23	12	15	82	99	79	100
123	23	12	20	61	72	36	77
126	23	12	20	145	86	116	85
128	23	12	15	122	80	125	82
131	23	12	15	128	94	109	95
137	23	12	20	71	103	58	98
139	23	12	10	58	98	82	98
140	23	12	20	76	97	56	94

Kuva 5. Esimerkki liikenneolosuhteiden keruuohjelman tuottamista LAM-tiedoista. Lyhenteistä q1 tarkoittaa suunnan 1 liikennemäärää ja v1 saman suunnan keskinopeutta. Suureet on laskettu viiden minuutin keskiarvoina. (Vilhunen 2002.)

Liikennetiedot siirrettiin liikennekeskukseen viiden minuutin välein, mistä ai-
heutui viiden minuutin keruuviive. Matka-aikatiedoilla oli myös matka-ajan
mittainen lisäviive, koska matka-aika voidaan mitata vasta, kun ajoneuvo
ohittaa jälkimmäisen mittauspisteen. Ennustemalleja tehtäessä oletettiin,
että tiedonsiirron aiheuttama viive oli matka-aikatietojen osalta nolla minuut-
tia ja LAM-tietojen osalta 10 minuuttia. Oletus ei ole matka-aikatietojen
osalta tällä hetkellä todenmukainen, mutta se on tilanne, johon tulee pyrkiä.

5.2 Aineisto

Ennustemallit tehtiin kesinä 2000 ja 2001 kerätyn aineiston perusteella. Matka-aika-aineisto oli järjestelmän tuottamaa raakadataa. Siinä oli mukana matkalla pysähtyneistä ja reitiltä poikenneista ajoneuvoista sekä rekisterikilpien väärin tulkinnasta johtuvia epärealistisia matka-aikahavaintoja.

Yksittäisiä matka-aikahavaintoja ei suodatettu. Tähän oli syynä se, että havaintoja kertyi viiden minuutin aikana liian vähän. Niistä viisiminuuttisista, joilta oli havaintoja, Kymijärvi-Heinola-suunnalla 33–60 prosentissa ja Heinola-Kymijärvi-suunnalla 60–78 prosentissa ajoneuvohavaintoja kertyi ainoastaan yksi tai kaksi (Taulukko 1). Silloin, kun suodatusta tehdään ajantasaisesti, havaintoja pitää olla riittävästi sen arviointiin, mitkä havainnoista ovat todenmukaisia ja mitkä eivät.

Taulukko 1. Ajoneuvohavaintojen määrän jakaumat (%) niissä viisiminuuttisissa, joiden ajalta on vähintään yksi havainto.

Linkki	Ajoneuvohavaintojen määrä, johon mediaani perustuu									
	1	2	3	4	5	6	7	8	9	≥ 10
Ky-He	36,4	23,9	16,4	9,8	6,0	3,5	1,9	1,0	0,5	0,7
Ky-Vi	25,0	19,3	13,9	11,2	8,7	6,4	4,4	3,4	2,3	5,3
Ky-Se	25,1	18,2	14,4	11,6	8,9	6,4	4,5	3,2	2,3	5,4
Se-He	26,1	19,7	15,2	11,6	8,6	6,2	4,0	2,9	2,0	3,8
Se-Vi	19,3	13,7	11,9	10,2	8,5	7,2	5,6	4,6	3,7	15,3
Vi-He	23,8	17,5	14,4	11,5	9,0	6,9	4,9	3,5	2,4	6,1
He-Ky	42,5	24,4	13,8	8,2	4,7	2,8	1,5	0,7	0,4	1,0
He-Se	53,6	24,4	11,8	5,5	2,2	1,1	0,6	0,3	0,2	0,4
He-Vi	36,3	24,5	15,1	8,9	5,4	3,0	2,1	1,4	0,7	2,5
Vi-Ky	36,9	23,0	15,2	8,8	5,3	3,4	2,3	1,6	0,9	2,7
Vi-Se	45,4	25,5	12,7	6,4	3,6	2,2	1,5	0,8	0,7	1,2
Se-Ky	49,1	25,4	12,7	5,6	3,1	1,5	0,9	0,5	0,3	0,9

6 ENNUSTEMALLIT

6.1 Syötesuureiden valinta

Tehtävänä oli siis ennustaa Kymijärvi–Heinola- ja Heinola–Kymijärvi-linkin matka-aika. Mallit antoivat vasteena tarkasteltavalle linkille seuraavan viiden minuutin aikana lähdössä olevien ajoneuvojen mediaanimatka-ajan. Vasteiden mediaaneja määritettäessä ajoneuvohavainnot oli järjestetty linkin alkupisteen ohitusajojen perusteella.

Se, ettei aineistosta suodatettu yksittäisiä poikkeavia matka-aikahavaintoja, rajasi syötemuuttujaehdokkaiden määrää. Yksittäinen poikkeava havainto vääristää esimerkiksi keskiarvoa ja keskihajontaa huomattavasti, joten tällaisia suureita ei kannattanut käyttää mallin syötteinä. Matka-aikasuureista mallin syöte-ehdokkaiksi valittiin mediaanimatka-ajat. Jos mediaani perustuu useaan havaintoon eikä poikkeavien havaintojen osuus ole suuri, mediaani on todenmukainen poikkeavista havainnoista huolimatta. Syötteinä olevat matka-aikamediaanit olivat kultakin linkiltä viimeksi mitattujen matka-aikojen mediaanit, eli niitä määritettäessä ajoneuvohavainnot järjestettiin linkin loppupisteen ohitusajan perusteella.

LAM-pisteissä määritetyistä pistekohtaisista suureista ajantasaisia tietoja saatiin ainoastaan liikennemääristä ja keskinopeuksista. Molemmat valittiin syötemuuttujaehdokkaiksi. Pistenopeuksien keskihajonta tai varianssi olisi myös ollut mielenkiintoinen syötesuure, sillä nopeuksien hajonnalla on taipumus kasvaa muutostilanteissa. Syötemuuttujaehdokkaita olivat siis eri osalinkkien matka-aikamediaanit sekä LAM-pisteissä mitatut liikennemäärät ja keskinopeudet.

Mallien syötesuureet valittiin ehdokkaiden joukosta sillä perusteella, kuinka hyvin ne korreloivat ennustettavan matka-ajan kanssa. Korrelaatiot määritettiin huomattavien ruuhkien aikaisista havainnoista. Sujuva liikenne jätettiin korrelaatiotarkastelun ulkopuolelle, koska sujuvan liikenteen ennustaminen ei ole yhtä haastavaa tai edes tärkeää kuin ruuhkien ennustaminen.

Matka-aikamediaanien korrelaatioista (Taulukko 2) havaittiin, että Kymijärvi–Heinola-mallin syötteistä Vierumäki–Heinola-linkin ja Heinola–Kymijärvi-mallin syötteistä muiden kuin Heinolasta alkavien linkkien matka-aikamediaanien korrelaatiot ennustettavan matka-ajan kanssa olivat erittäin alhaisia. Tähän oli syynä se, ettei kyseisillä osalinkeillä ollut opetusaineiston keruujaksona lainkaan ruuhkaa (Innamaa ym. 2002) eikä niiden matka-ajassa näin ollen tapahtunut muutoksia koko linkin matka-ajan muuttuessa.

Taulukko 2. Korrelaatio ennustettavien matka-aikojen ja syöte-ehdokkaidenvälillä.

Syöte-ehdokas		Ky–He	Syöte-ehdokas		He–Ky
Matka-aika	Ky–He	0,76	Matka-aika	He–Ky	0,67
	Ky–Vi	0,83		He–Se	0,75
	Ky–Se	0,72		He–Vi	0,82
	Se–He	0,67		Vi–Ky	0,00
	Se–Vi	0,74		Vi–Se	-0,05
	Vi–He	0,19		Se–Ky	0,13
LAM 601	Liikennemäärä	-0,09	LAM 601	Liikennemäärä	0,21
	Keskinopeus	-0,22		Keskinopeus	0,07
LAM 424	Liikennemäärä	0,46	LAM 628	Liikennemäärä	0,31
	Keskinopeus	-0,14		Keskinopeus	-0,07

Syötteiksi päätettiin valita kaikki ne suureet, joiden korrelaatio ennustettavan suureen suhteen oli vähintään 0,2 (Taulukko 2). Tällöin Kymijärvi–Heinola-mallin syötesuureiksi valittiin LAM-pisteen 601 keskinopeus, LAM-pisteen 424 liikennemäärä ja kaikkien muiden suunnan osalinkkien paitsi Vierumäki–Heinola-linkin mediaanimatka-ajat. Heinola–Kymijärvi-mallin syötteiksi valittiin ainoastaan Heinolasta alkavien linkkien matka-ajat sekä LAM-pisteiden 601 ja 628 liikennemäärät. Muiden suureiden korrelaatio ennustettavien matka-aikojen kanssa oli alhainen.

6.2 Mallien rakentaminen

Malleja tehtäessä kerätystä aineistosta käytettiin kello 8:00–22:00 välisiä havaintoja. Syöteaineisto koottiin 3 x 5 minuutin aikasarjoista, eli kustakin syötesuureesta annettiin kolme viimeistä arvoa.

Koska matka-aikahavaintoja ei saatu jokaiselta viiden minuutin jaksolta jokaiselta osalinkiltä, täytyi matka-aikasuureille kehittää päivityssääntö, jotta malli voisi toimia mahdollisimman häiriöttä pienistä katkoksista huolimatta. Matka-aikasyötteen arvo päätettiin pidetään samana, kunnes saatiin uusi mittaustulos. Mikäli tämä uusi mediaani perustui useamman kuin yhden ajoneuvon matka-aikaan, syötesuureen arvo päivitettiin aina. Kahden havainnon mediaaniksi määritettiin matka-ajoista lyhyempi. Jos mediaani perustui yhden ajoneuvon matka-aikaan, syötesuureen arvo päivitettiin ainoastaan siinä tapauksessa, ettei uusi arvo eronnut vanhasta yli 20 prosenttia. Tällä haluttiin varmistaa se, ettei yksittäinen poikkeava havainto päässyt vaikuttamaan matka-aikasyötteen arvoon liikaa. Lopullisen mallin mahdollisimman keskeytyksetöntä toimintaa varten varauduttiin yksittäisten ilmaisimien rikkoutumiseen ja opetettiin neuroverkot, jotka toimivat ilman kyseiseltä ilmaisimelta saatavia syötetietoja.

Mallien neuroverkot opetettiin kahdenlaisella opetusaineistolla: toinen koostui kaikkien aineistossa olevien päivien havainnoista ja toinen oli koottu pelkästään ruuhkapäivien havainnoista, koska tällaisella opetusjoukolla opetettulla mallilla on aikaisemmin (Innamaa ym. 2002) saavutettu parempi ruuhkan ennustamiskyky kuin koko aineistolla opetetulla mallilla. Päivä määriteltiin tässä yhteydessä ruuhkapäiväksi, mikäli matka-aika oli sen aikana nousnut vähintään 30 minuutin ajaksi yli 25 minuuttiin. Kymijärvi–Heinola-välillä näitä päiviä oli 14 ja Heinola–Kymijärvi-välillä viisi. Syöteaineistoista rajattiin pois ne jaksot, jolloin kahden ilmaisun välinen aika ylitti 30 minuuttia jollakin osalinkillä tai jossakin LAM-pisteessä.

Aineisto esikäsiteltiin normeeraamalla se siten, että kunkin syötemuuttujan arvojen keskiarvoksi tuli nolla ja keskihajonnaksi yksi. Piilokerroksen aktivaatiofunktioiksi valittiin hyperbolinen tangentti ja vastekerroksen aktivaatiofunktioiksi lineaarinen funktio.

Mallien testaaminen suoritettiin kaikella olemassa olevalla aineistolla. Testiaineisto oli siis sama kaikille malleille opetusaineistosta riippumatta.

7 ENNUSTEMALLIEN HYVYYS

Innamaa ym. (2002) totesivat ruuhkassa oikein ennustettujen matka-aikojen osuuden parhaaksi mallien hyvyyden mittariksi. Myös tässä tutkimuksessa kyseistä osuutta käytettiin mallien vertailussa pääkriteerinä. Muina kriteereinä käytettiin ruuhkassa oikein ennustettujen matka-aikojen osuutta sekä keskimääräistä neliövirhettä, virheen itseisarvoa ja suhteellisen virheen itseisarvoa.

Erilaisilla aineistoilla opetettujen mallien tulokset on koottu taulukkoon 3. Kymijärvi–Heinola-linkillä koko aineistolla opetetut mallit olivat kaikilla kriteereillä mitattuna hieman parempia kuin vastaavat pelkkien ruuhkapäivien aineistolla opetetut mallit. Heinola–Kymijärvi-linkillä suunta oli sama, mutta ero ei ollut yhtä selvä. Ruuhkassa oikein ennustettujen matka-aikojen osuudella mitattuna niistä malleista, joiden syötteissä ei ollut LAM-tietoja, pelkkien ruuhkapäivien aineistolla opetettu malli oli jopa parempi kuin koko aineistolla opetettu malli. Molemmilla suunnilla ne mallit, joiden syötteissä LAM-suureet olivat mukana, olivat parempia kuin ne mallit, joiden syötteet koostuivat pelkistä matka-aikamediaaneista.

Taulukko 3. Erilaisten ennustemallien tulokset. O% tarkoittaa oikein ennustettujen matka-aikojen osuutta, RO% ruuhkassa oikein ennustettujen matka-aikojen osuutta, MSE on keskimääräinen neliövirhe, MAE keskimääräinen virheen itseisarvo ja MARE keskimääräinen suhteellisen virheen itseisarvo.

Linkki	Opetus- aineisto	LAM- syötteet	O%	RO%	MSE	MAE	MARE
Ky–He	Ruuhka- päivät	ei	96 %	67 %	3,5	1,2	6 %
		mukana	97 %	72 %	3,1	1,1	6 %
	Koko aineisto	ei	97 %	67 %	2,6	1,0	5 %
		mukana	98 %	73 %	2,2	0,9	5 %
He–Ky	Ruuhka- päivät	ei	99 %	82 %	2,0	1,0	6 %
		mukana	99 %	84 %	1,8	0,9	5 %
	Koko aineisto	ei	99 %	80 %	1,6	0,9	5 %
		mukana	99 %	85 %	1,6	0,9	5 %

Ennustemallin täytyy varautua myös tilanteisiin, joissa jokin ilmaisimien ei toimi. Jos katkos on korkeintaan puolen tunnin mittainen, voidaan syötesuureen arvon olettaa pysyvän samana, kunnes saadaan uusi arvo. Jos katkos on pidempi, ennuste tehdään mallilla, joka on opetettu aineistolla, josta kyseisen ilmaisimen syötesuureet puuttuvat. Puutteellisilla syötetiedoilla opetettujen mallien opetusjoukko koottiin koko aineistosta.

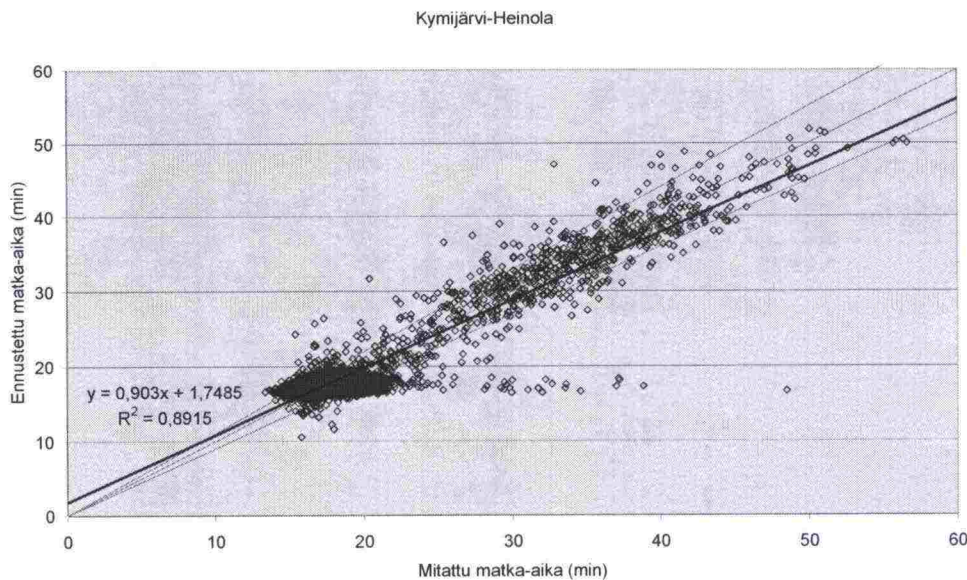
Matka-aikatiedotusjärjestelmän uskottavuuden kannalta on tärkeää, ettei liian epävarmoja ennusteita esitetä. Tästä syystä liian monen tai ennusteen kannalta kriittisen ilmaisimen rikkoutuessa ennustetta ei tehdä. Puutteellisilla syötetiedoilla opetettujen mallien tulokset on esitetty taulukossa 4. Taulukossa mainittujen tapausten lisäksi tilanteissa, joissa useampi kuin yksi kamera on poissa toiminnasta, ennustetta ei tehdä.

Taulukko 4. Vajaalla syötteellä opetettujen ennustemallien tulokset. O% tarkoittaa oikein ennustettujen matka-aikojen osuutta, RO% ruuhkassa oikein ennustettujen matka-aikojen osuutta, MSE on keskimääräinen neliövirhe, MAE keskimääräinen virheen itseisarvo ja MARE keskimääräinen suhteellisen virheen itseisarvo.

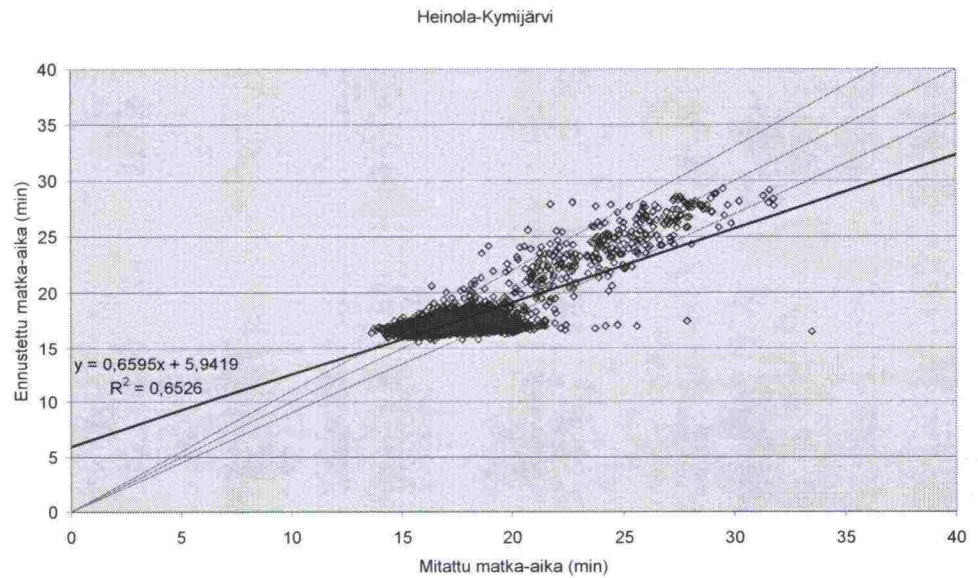
Linkki	Toimimaton kamera	Toimimaton LAM-ilmaisoin	O%	RO%	MSE	MAE	MARE
Ky-He	-	-	98 %	73 %	2,2	0,9	5 %
		424	97 %	66 %	2,5	1,0	5 %
		601	98 %	70 %	2,3	0,9	5 %
		424 ja 601	97 %	67 %	2,6	1,0	5 %
	Ky	EI ENNUSTETTA					
	Se	-	97 %	65 %	2,5	1,0	5 %
		424	EI ENNUSTETTA				
		601	97 %	63 %	2,6	1,0	5 %
		424 ja 601	EI ENNUSTETTA				
	Vi	-	97 %	66 %	2,6	1,0	5 %
		424	EI ENNUSTETTA				
		601	97 %	63 %	2,6	1,0	5 %
		424 ja 601	EI ENNUSTETTA				
	He	-	98 %	71 %	2,3	1,0	5 %
		424	EI ENNUSTETTA				
		601	97 %	67 %	2,4	1,0	5 %
		424 ja 601	EI ENNUSTETTA				
He-Ky	-	-	99 %	85 %	1,6	0,9	5 %
		601	99 %	80 %	1,6	0,9	5 %
		628	99 %	83 %	1,6	0,9	5 %
		601 ja 628	99 %	80 %	1,6	0,9	5 %
	He	EI ENNUSTETTA					
	Vi	-	99 %	78 %	1,7	0,9	5 %
		601	98 %	76 %	1,7	0,9	5 %
		628	98 %	71 %	1,8	0,9	5 %
		601 ja 628	EI ENNUSTETTA				
	Se	-	99 %	84 %	1,6	0,9	5 %
		601	99 %	81 %	1,6	0,9	5 %
		628	99 %	82 %	1,7	0,9	5 %
		601 ja 628	EI ENNUSTETTA				
	Ky	-	99 %	81 %	1,6	0,9	5 %
		601	99 %	80 %	1,6	0,9	5 %
		628	99 %	82 %	1,7	0,9	5 %
		601 ja 628	EI ENNUSTETTA				

Taulukon 3 malleja verrattiin parhaisiin aikaisempiin malleihin. Aikaisemmin tehdyistä malleista (Innamaa ym. 2002) paras Kymijärvi-Heinola-linkin malli oli oikeassa 73 prosenttia ajasta ruuhkassa ja Heinola-Kymijärvi-linkin malli 84 prosenttia ajasta. Tulokset ovat yhtenevät nyt tehtyjen mallien kanssa (Taulukko 3). Tosin malleja ei voi verrata aivan täysin, sillä Innamaan ym. (2002) tutkimuksessa mallit ennustivat seuraavan minuutin aikana linkille lähtevien ajoneuvojen keskimatka-aikaa ja nyt tehtyt mallit seuraavan viiden minuutin aikana linkille lähtevien ajoneuvojen mediaanimatka-aikaa. Tulokseen voi joka tapauksessa olla tyytyväinen, sillä Innamaan ym. (2002) malleilla oli monipuolisemmat syötteet kuin nyt tehdyillä malleilla.

Molemmille linkeille mitatut ja kaikkien syötteiden avulla ennustetut mediaanimatka-ajat on esitetty kuvissa 6 ja 7. Molemmissa kuvissa on havaittavissa joukko havaintoja, joissa on mitattu yli 22 minuutin matka-aikoja, mutta ennuste on pysynyt alle 20 minuutissa. Esimerkiksi Kymijärvi-Heinola-linkillä tällaisia tapauksia oli 46, joista 40 oli tilanteita, joissa mitatuissa arvoissa oli yksittäinen piikki. Loput kuusi tapausta olivat joko lieviä, lyhytkestoisia ruuhkia tai pidempien ruuhkien alkuja, joista ennustemalli myöhästyi. Suurinta osaa näistä "virheistä" ei siis voi pitää vakavina.



Kuva 6. Kymijärvi-Heinola-mallin ennustamat ja linkille mitatut mediaanimatka-ajat, havaintopisteisiin sovitettu regressiosuora (paksumpi, tumma suora), sen sovitussaste (R^2) sekä suorat $y = x$ ja $y = x \pm 10\%$ (ohuet, vaaleammat suorat).



Kuva 7. Heinola–Kymijärvi-mallin ennustamat ja linkille mitatut mediaanimatka-ajat, havaintopisteisiin sovitettu regressiosuora (paksumpi, tumma suora), sen sovitussaste (R^2) sekä suorat $y = x$ ja $y = x \pm 10\%$ (ohuet, vaaleammat suorat).

Liitteessä 1 on esitetty yksittäisistä ruuhkapäivistä, kuinka ennustettu mediaanimatka-aika käyttäytyy suhteessa mitattuihin arvoihin. Lisäksi liitteessä on kuvaajat, joissa on esitetty mitattujen syötesuureiden kehitys samassa kuvassa ennustettavan (= mitatun) mediaanimatka-ajan kanssa. Liitteen kuvista on havaittavissa, että kun linkkiä edeltävän LAM-pisteen (Kymijärvi–Heinola-linkillä LAM 424 Lahden eteläpuolella ja Heinola–Kymijärvi-linkillä LAM 628 Heinolan pohjoispuolella) liikennemäärä on noussut raja-arvon 100 ajon./h yli ennen kuin matka-aika alkaa kasvaa, malli ei yleensä myöhästy ruuhkan alun ennustamisesta. Sitä vastoin, jos kyseinen liikennemäärä on alhainen ruuhkan alkaessa, malli myöhästy.

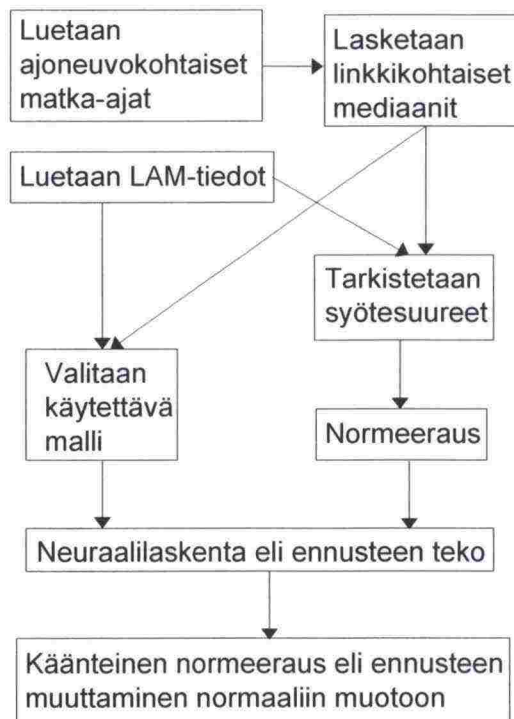
8 MALLIEN OHJELMOIMINEN

Mallit tehtiin Matlab-ohjelmistolla ja ne piti ohjelmoida yleiseen muotoon, ennen kuin ne voitiin asentaa Tiehallinnon koneille. Ohjelmointi suoritettiin LabVIEW-ohjelmointikielellä.

Ennustemalliohjelman toimintaperiaate on esitetty kuvassa 8. Ennusteen teko alkaa matka-aikahavaintojen (Kuva 4) ja LAM-tietojen (Kuva 5) lukemisella. Tämä tapahtuu 5 minuutin välein. Matka-ajoista määritetään linkkikohtaiset mediaanit siten, että jos havaintoja on kahdesta ajoneuvosta, mediaaniksi määritetään matka-ajoista lyhyempi. Syötesuureiden arvot päivitetään luvun 6.2 mukaisesti ja saadut arvot normeerataan. Jos syötesuuretta merkitään symbolilla x , normeerattu syötesuure x_n saadaan seuraavan yhtälön avulla.

$$x_n = \frac{x - m_{in}}{s_{in}}$$

Yhtälössä m_{in} on opetusaineiston avulla määritetty syötesuureen odotusarvo ja s_{in} keskihajonta.



Kuva 8. Ennustemallin toimintaperiaate.

Jos jonkin ilmaisimen viimeisestä havainnosta on pitkä aika (yli 30 minuuttia), valitaan malli, joka ei käytä tämän ilmaisimen tietoja hyväkseen (Taulukot 4 ja 5). Ennustetta ei tehdä lainkaan Taulukko 4 mainittujen erityistapausten lisäksi tilanteissa, joissa useampi kuin yksi kamerailmaisim on poissa toiminnasta.

Taulukko 5. Puutteellisien matka-aikatietojen perusteella mallin valinnan yhteydessä tehtävä tulkinta ilmaisimien toimintakyvystä. Muut puuttuvien matka-aikatietojen yhdistelmät johtavat tilanteeseen, jossa ennustetta ei kannata tehdä.

Kymijärvi–Heinola-malli		Heinola–Kymijärvi-malli	
Puuttuvat matka-aikatiedot	Toimimattoman ilmaisimen sijainti	Puuttuvat matka-aikatiedot	Toimimattoman ilmaisimen sijainti
Ky–He	Heinola	He–Ky	Kymijärvi
Ky–Vi	Vierumäki	He–Se	Seesta
Ky–Se	Seesta	He–Vi	Vierumäki
Se–He	Heinola		
Se–Vi	Vierumäki		
Ky–He, Se–He	Heinola		
Ky–Vi, Se–Vi	Vierumäki		
Ky–Se, Se–He	Seesta		
Ky–Se, Se–Vi	Seesta		
Se–He, Se–Vi	Seesta		
Ky–Se, Se–He, Se–Vi	Seesta		

Ennuste tehdään valitulla neuroverkkomallilla ja saatu vaste palautetaan normaaliin muotoon (minuuteiksi). Matemaattisesti ennusteen teko tapahtuu seuraavasti.

$$\mathbf{y}_n = \mathbf{f}_2(\mathbf{W}_{21}\mathbf{f}_1(\mathbf{W}_{11}\mathbf{x}_n + \mathbf{b}_1) + \mathbf{b}_2)$$

Yhtälössä \mathbf{y}_n on mallin antama vaste. Matriisissa \mathbf{W}_{11} on syöte- ja piilokerroksen välisten summien painokertoimet ja matriisissa \mathbf{W}_{21} vastaavasti piilo- ja vastekerroksen välisten summien painokertoimet. Syötteet on koottu matriisiin \mathbf{x}_n , ja \mathbf{b}_1 ja \mathbf{b}_2 ovat harhakertoimia, jotka mahdollistavat sen, ettei malli rajoitu origon kautta kulkeviin tasoihin. Funktio \mathbf{f}_1 on piilokerroksen aktivaatiofunktio, joka on muotoa

$$f_1(z) = \frac{2}{1 + e^{-2z}} - 1$$

ja \mathbf{f}_2 on vastekerroksen aktivaatiofunktio, joka on muotoa

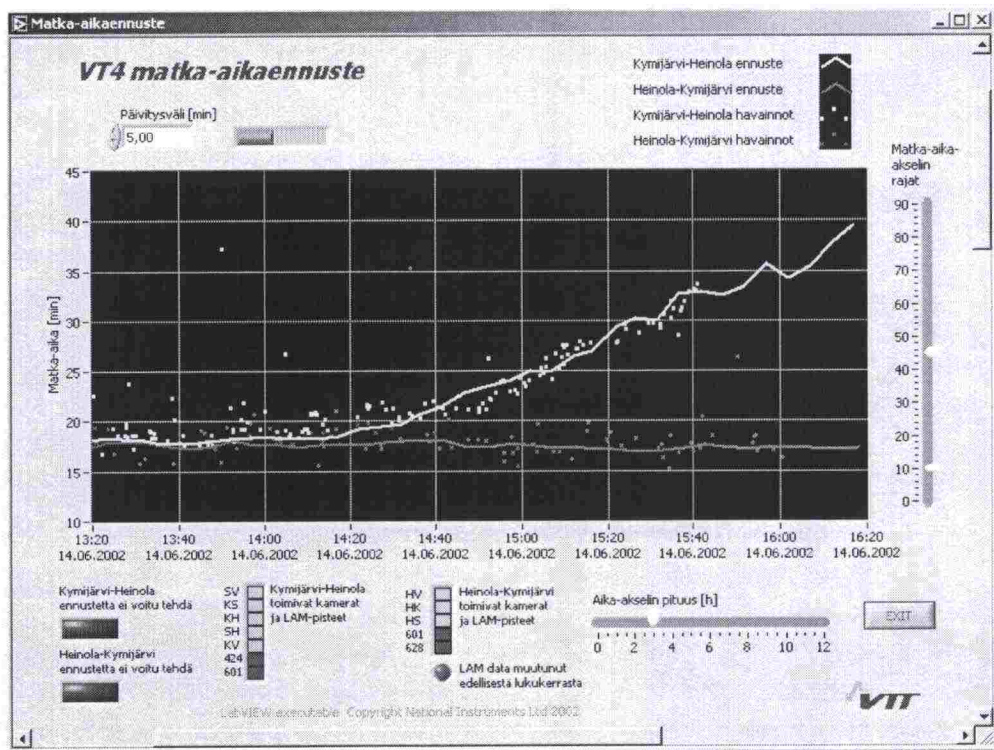
$$f_2(z) = z$$

Vaste (y_n) saadaan palautettua normaaliin muotoon (y) seuraavan yhtälön avulla.

$$y = s_{out} \cdot y_n + m_{out}$$

Yhtälössä m_{out} on opetusaineiston avulla määritetty vastesuureen odotusarvo ja s_{out} keskihajonta.

Ennustettu matka-aika esitetään tietokoneen näytöllä yhdessä toteutuneiden matka-aikojen kanssa suunnittain (Kuva 9). Toteutuneiden matka-aikahavaintojen ajankohta (x-akselilla) on se hetki, jolloin he ohittavat ensimmäisen kamerapisteen eli suunnilleen se hetki, jolloin he saavat muuttuvan opasteen tiedon odotettavissa olevasta matka-ajasta. Vastaavaksi ennusteen esitysajankohdaksi on valittu hetki, jolloin viimeisin matka-aika on mitattu. Todellisuudessa ennuste tehdään tiedonsiirtoviiveen verran myöhemmin, mutta esitetty tilanne on ihanne, johon tulisi pyrkiä. Se on toisin sanoen aikaisin hetki, jolloin kyseinen ennuste voidaan tehdä.



Kuva 9. Ennustemallin käyttöliittymän ensimmäinen näyttö, jossa ennustettu matka-aika on esitetty suunnittain viivalla ja toteutuneet matka-ajat pisteinä. Näytön vasemmassa alanurkassa syttyvät varoitusvalot, jos ennustetta ei voida tehdä. Pienemmät liikennevalot kertovat suunnittain, mitä ilmaisimia mallilla on käytettävissään. Vasemman ylänurkan palkki kertoo, kuinka paljon aikaa on jäljellä seuraavan ennusteen tekohetkeen.

Käyttöliittymän ensimmäisen näytön (Kuva 9) alalaidasta nähdään, minkä LAM-ilmaisimien tai linkkivälien tietoihin ennuste voidaan perustaa. Vihreä valo kertoo, että viimeisin tieto on alle puolen tunnin ikäistä, ja punainen valo, että tieto on liian vanhaa ennusteen käytettäväksi. Vasempaan alakulmaan syttyvät punaiset valot, jos ennustetta ei voida tehdä.

Käyttöliittymän toisella näytöllä (Kuva 10) on tietoa havaintojen lukumääristä sekä viimeksi laskettujen matka-aikamediaanien arvot. Käyttöliittymän kolmannelta näytöltä (Kuva 11) voi selata viimeksi luettuja liikennetietoja.

Matka-aikaennuste					
	Seuraavaksi luettavien ID:t	Alle puoli tuntia vanhat havainnot		Viimeksi luettujen havaintojen mediaanit ja lukumäärät	
0 HV Heinola-Vierumäki	5315444	8		6,11	4
1 VH Vierumäki-Heinola	5315445	12		7,30	1
2 VS Vierumäki-Seesta	5315448	18		5,27	3
3 SV Seesta-Vierumäki	5315458	47		13,90	10
4 SK Seesta-Kymijärvi	5315460	23		5,58	2
5 KS Kymijärvi-Seesta	5315477	77		11,58	17
6 HK Heinola-Kymijärvi	5315480	19		17,00	3
7 KH Kymijärvi-Heinola	5315486	29		32,96	6
8 HS Heinola-Seesta	5315488	22		11,03	2
9 SH Seesta-Heinola	5315496	40		19,67	8
10 VK Vierumäki-Kymijärvi	5315502	32		11,15	6
11 KV Kymijärvi-Vierumäki	5315507	40		27,27	5

Ilmaisin oletetaan toimivaksi, jos siltä on saatu dataa, joka ei ole puolta tuntia vanhempaa.

Kuva 10. Ennustemallin käyttöliittymän toinen näyttö, josta näkyvät muun muassa viimeisen hakukerran havaintojen lukumäärät linkeittäin ja näiden perusteella lasketut matka-aikamediaanit sekä alle puolen tunnin ikäisten havaintojen lukumäärät.

Matka-aikaennuste

Näyttää lukuhetkellä tulevaa dataa

5315502	11	FI:XXXXXX	1644	2002-06-14 15:45:53	2002-06-14 16:13:17	7	3 26713481	26715178
5315503	11	FI:XXXXXX	1638	2002-06-14 15:45:15	2002-06-14 16:12:33	7	3 26713475	26715169
5315504	11	FI:XXXXXX	1411	2002-06-14 15:49:41	2002-06-14 16:13:12	7	3 26713795	26715177
5315505	11	FI:XXXXXX	1624	2002-06-14 15:49:14	2002-06-14 16:16:18	7	3 26713791	26715198
5315506	11	FI:XXXXXX	1636	2002-06-14 15:45:50	2002-06-14 16:13:06	7	3 26713480	26715176

Matka-akadata

Linkkiväli	0	5314741	0	350	2002	6	14	15	19
Rmi	0	5314817	0	370	2002	6	14	15	22
Sarake	0	5314741	0	350	2002	6	14	15	19
		5314866	0	393	2002	6	14	15	27
		5314867	0	421	2002	6	14	15	26
		5314741	0	350	2002	6	14	15	19

LAM data

Rmi	0	1424	0	0	0	0	0	0	0
Sarake	0	1601	0	0	0	0	0	0	0
		2601	0	0	0	0	0	0	0
		2628	2002	6	14	9	20	3106880400	3106905875 254

Kuva 11. Ennustemallin käyttöliittymän kolmas näyttö, jonka avulla nähdään viimeksi luetut liikennetiedot.

Ennusteohjelma tekee kolmenlaisia lokitiedostoja. Yhteen se kokoaa tiedot tiedon lukuhetkistä, valitusta mallista, tehdyistä ennusteista sekä syötesuurten arvoista. Kahdessa muussa lokitiedostossa on matka-aika- sekä LAM-tiedot sellaisina kuin ohjelma on ne lukenut. Luetut liikennetiedot ja ensimmäisen lokitiedoston mallitiedot voidaan yhdistää juoksevan tietojenlukkerrasta kertovan indeksin avulla.

9 JOHTOPÄÄTÖKSIÄ JA SUOSITUKSIA

Kymijärvi–Heinola-malli ennusti ruuhkassa oikein 73 prosenttia ja Heinola–Kymijärvi-malli 85 prosenttia ajasta. Malleja voidaan pitää varsin hyvinä – etenkin, kun ottaa huomioon, ettei havaintoja voitu suodattaa ja syötesuureet olivat sängen rajoitetut.

Kymijärvi–Heinola-suunnalla voitiin varautua kymmeneen ja Heinola–Kymijärvi-suunnalla 13:een ilmaisimien toiminnan kannalta erilaiseen tilanteeseen, joissa malli antoi tyydyttäviä tuloksia. Jäljelle jäi tilanteita, joissa riittävän hyvää ennustetta ei voida tehdä, mutta malleja voidaan pitää varsin hyvin varautuneina yksittäisten ilmaisimien rikkoutumiseen.

Miten malleja voisi parantaa? Havaittiin, kuinka tärkeää ensimmäistä kamerapistettä edeltävän LAM-pisteen liikennemäärätieto on mallille. Ennuste vastasi erittäin hyvin ruuhkaa myös ruuhkan alussa, kun linkkiä edeltävän LAM-pisteen liikennemäärä oli kohonnut ennen matka-ajan kasvua. Silloin, kun matka-aika kasvoi ennen kuin linkkiä edeltävän LAM-pisteen liikennemäärä alkoi kasvaa, ennuste myöhästyi ruuhkan alusta selvästi.

Nykyisellään linkkiä edeltävät LAM-pisteet sijaitsevat Lahden ja Heinolan toisella puolella, jolloin näiden LAM-pisteiden liikennemäärissä ovat mukana Lahteen ja Heinolaan menevät ajoneuvot eikä näistä kaupungeista tulevista ajoneuvoista ole mitään tietoa. Suhteellisen alhaisesta korrelaatiosta huolimatta, liikennemäärätiedon merkitys oli suuri. Tilanne paranisi varmasti vielä huomattavasti, jos LAM-pisteet olisi sijoitettu ensimmäisen kamerapisteen ja Lahden tai Heinolan väliin, jolloin saataisiin tieto linkeille tulossa olevista ajoneuvoista.

Jos saapuvien liikennevirtojen lisäksi tunnettaisiin poistuvat liikennevirrat, voitaisiin muusta kuin ylikysynnästä aiheutuvat satunnaiset liikenteen häiriöt havaita nopeammin ja matka-aikaennuste olisi tarkempi. Käytännössä tällaiset tilanteet näkyisivät sisään menevien ja ulos tulevien liikennemäärien erotuksen kasvuna (vertaa häiriöiden havaitsemisalgoritmit).

Toinen tärkeä asia, jossa liikennemäärätieto voisi auttaa, on sen arvioiminen, kuvaavatko mitatut matka-ajat osuuden todellista liikennetilannetta. Jos liikennemäärästä ei ole mitään tietoa, on vaikea sanoa, johtuvatko pitkät matka-ajat mahdollisesti ylikysynnästä, satunnaisesta häiriöstä vai kellojen kalibroinnin pettämisestä. Kysymys on ajankohtainen tilanteissa, jolloin matka-aikahavaintoja saadaan vähän esimerkiksi suoraan kameraan paistavan auringon takia.

Ihanteellista olisi, jos liikennemäärätiedot saataisiin suoraan matka-aikojen mittausjärjestelmästä. Näin välttyttäisiin päällekkäisiltä järjestelmiltä. Tämä tosin vaatisi järjestelmän parantamista siten, että suurempi osa ajoneuvoista tulisi havaituksi ja että kaikkia kaistoja seurattaisiin. Myös kustannukset kasvaisivat, sillä suurempaan otoskokoön pääsevät järjestelmät ovat huomattavasti nykyistä kalliimpia.

Pieni otoskoko ei ole ongelma, jos halutaan estimoida keskimääräinen tai mediaanimatka-aika. Liikennemääräestimaattiin otoskoko ei kuitenkaan riitä. Samoin matka-ajan hajonnan luotettavaan arviointiin otoskoko on liian pieni. Suuri hajonta voisi olla merkinä häiriöistä tai muutoksista liikennetilanteessa.

Olisikin tärkeää, että kamerat asennettaisiin kaikille kaistoille, jos halutaan tarkka kuva matka-ajasta muulloinkin kuin ruuhkassa, jolloin ohituskaistan merkitys pienenee. Toisaalta tienkäyttäjille juuri ruuhkatiedot ovat tärkeimpiä.

LÄHDELUETTELO

- DEMUTH H, BEALE M (1998). *Neural Networks Toolbox for Use with Matlab*. User's Guide, Version 3. The Math Works Inc. s. 5-1 – 5-58.
- ELORANTA T (1999). *Rekisterikilpien tunnistukseen perustuva liikenteen automaattinen matkanopeuden seuranta*. Tielaitoksen selvityksiä 46/1999. Uudenmaan tiepiiri, Tielaitos, Helsinki. 149 s.
- INNAMAA S, LANNE L, VANHANEN K, PURSULA M (2002). *Pääteiden lyhyen aikavälin matka-aikaennusteet*. Tiehallinnon selvityksiä 5/2002. Tiehallinto, Helsinki. 81 + 16 s.
- INNAMAA S, PURSULA M (2000). *Liikennemäärän ja nopeuden lyhyen aikavälin ennustaminen*. Tielaitoksen selvityksiä 54/2000. Tielaitos, Helsinki. 101 + 3 s.
- KILJUNEN M, SUMMALA H (1996). *Ruuhkaisuuden kokeminen ja liikennetiedottaminen*. Tielaitoksen selvityksiä 25/1996, Tielaitos, Keskushallinto, Helsinki. 77 s.
- KOKKINEN M (2002). Matti Kokkisen, Simulus Oy haastattelu sähköpostin välityksellä 15.1.2002.
- LEE S, KIM D, KIM J, CHO B (1998). *Comparison of Models for Predicting Short-Term Travel Speeds*. Conference CD-ROM, 5th World Congress on Intelligent Transport Systems, 12 – 16 October 1998, Seoul, Korea. 9 s.
- LUOMA S (1998). *Tieliikenteen sujuvuus ja sen mittaaminen*. Diplomityö, Teknillinen korkeakoulu, Liikennetekniikka. 122 s.
- SMITH B, DEMETSKY M (1994). *Short-Term Traffic Flow Prediction: Neural Network Approach*. Transportation Research Record 1453. S. 98 - 104.
- SMITH B, DEMETSKY M (1997). *Traffic Flow Forecasting: Comparison of Modeling Approaches*. Journal of Transportation Engineering, Vol. 123, No. 4, July / August 1997. S. 261 - 266.
- TIELAITOS (1999). *Liikenteen automaattinen mittaus 1998*. Tielaitoksen sisäisiä julkaisuja 52/1999, Tielaitos, Helsinki. 240 s.
- VILHUNEN J (2002). Tieinsinööri Jyri Vilhusen, Tiehallinto haastattelu sähköpostin välityksellä 23. – 25.1.2002.

LIITTEET

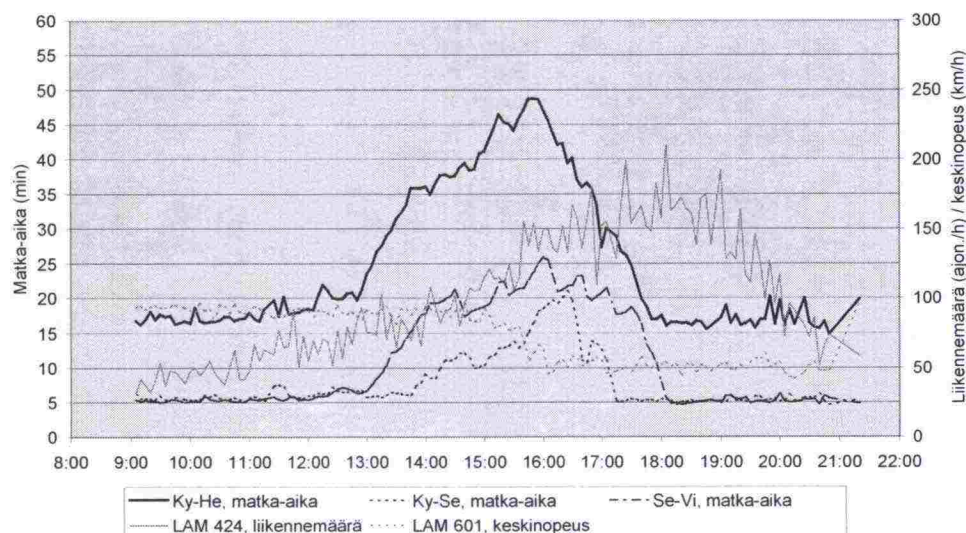
Liite 1: Ruuhkapäivien ennusteet

RUUHKAPÄIVIEN ENNUSTEET

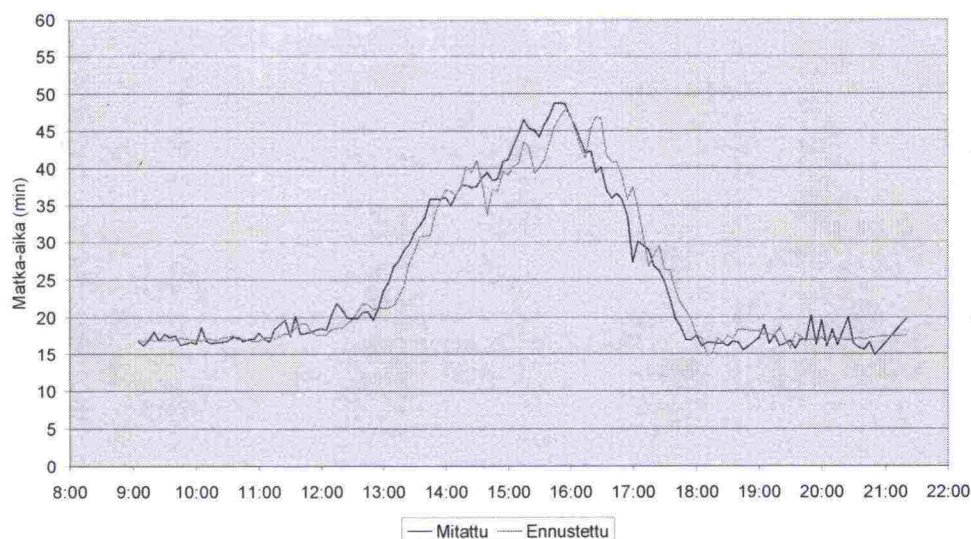
Kullakin sivulla on ylemmässä kuvassa ennustettava suure (vahva viiva) sekä vastaavalla hetkellä mitatut syötesuureet ja alemmassa kuvassa ennuste sekä vastaava mitattu arvo. Ennusteet on tehty kaikkien päivien havainnoista kootuilla opetusaineistoilla opetettuihin malleihin, jotka saivat syötteikseen matka-aikamediaanien lisäksi LAM-tiedot.

Kymijärvi-Heinola

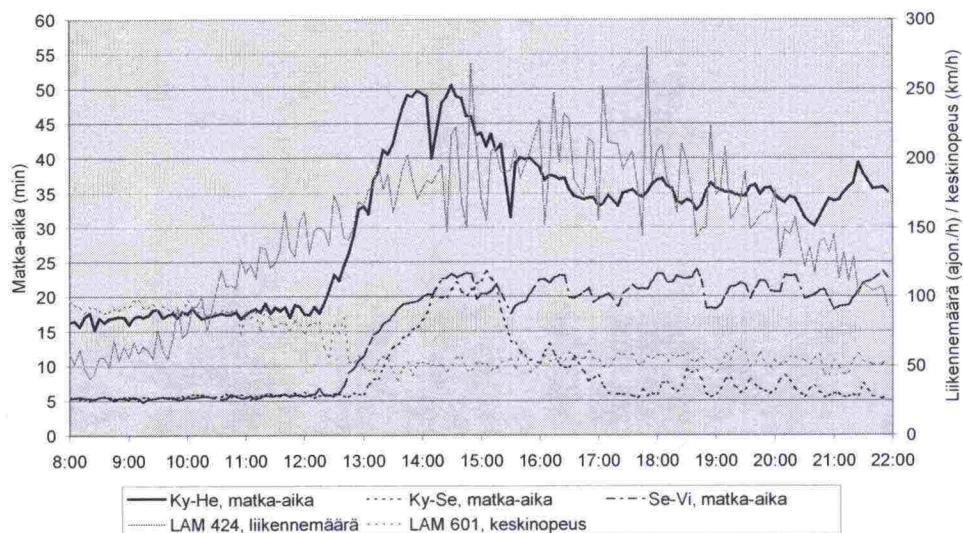
Kymijärvi-Heinola, 30.6.2000



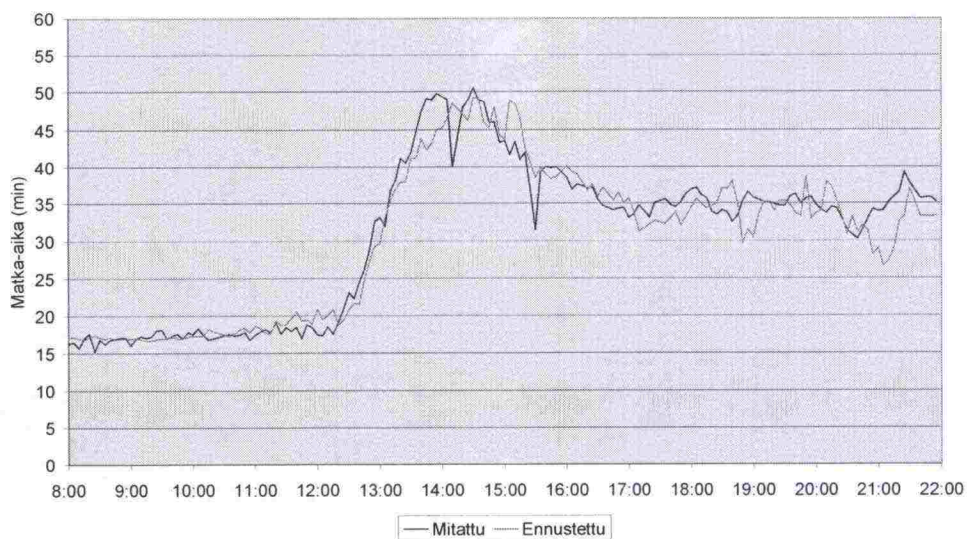
Kymijärvi-Heinola, 30.6.2000



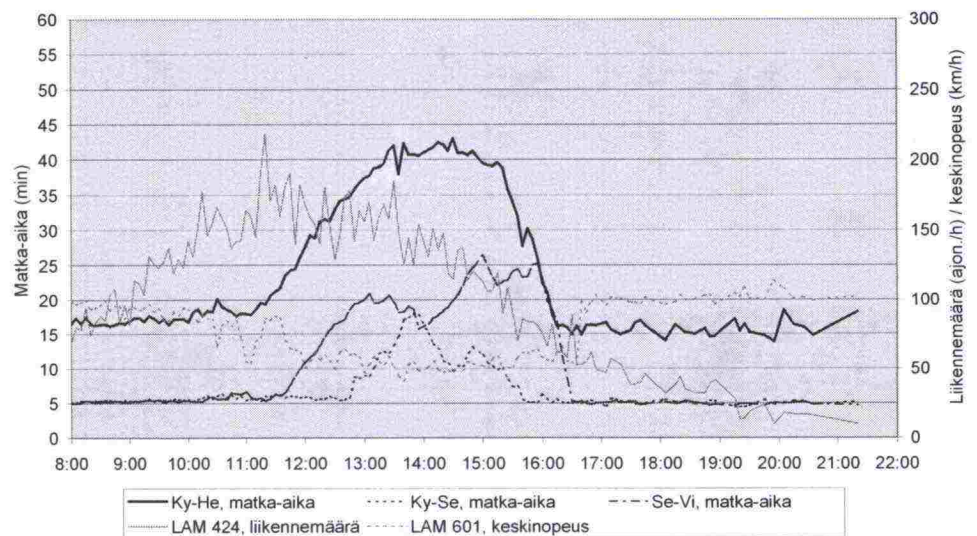
Kymijärvi-Heinola, 21.6.2001



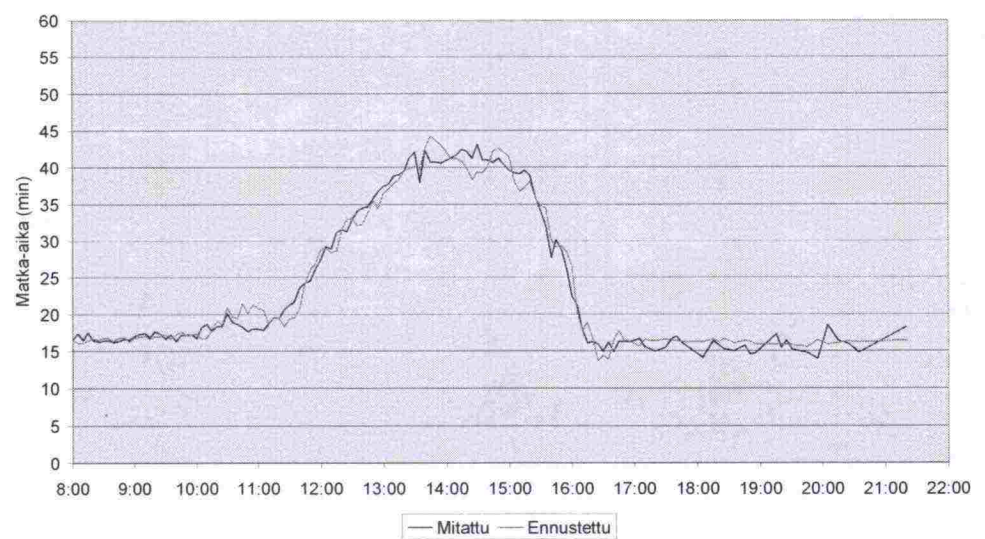
Kymijärvi-Heinola, 21.6.2001



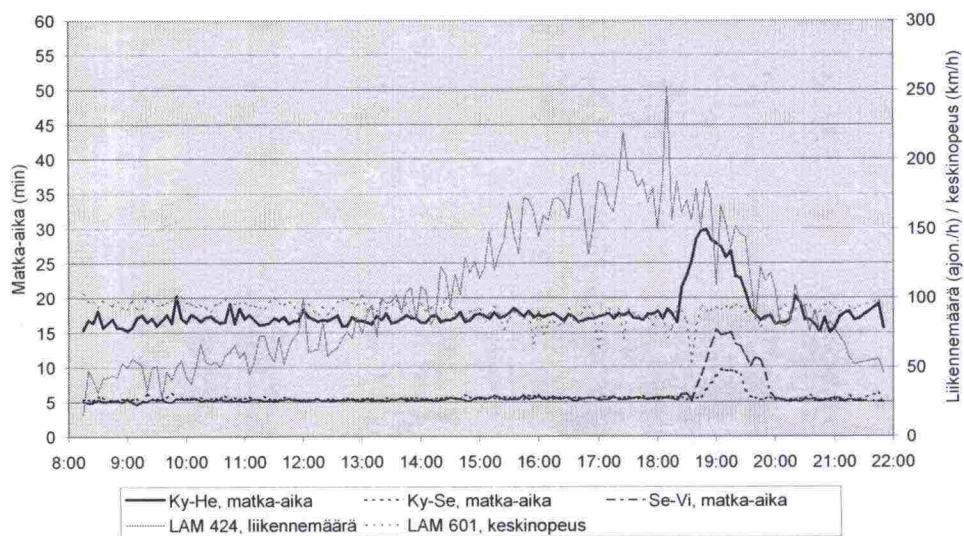
Kymijärvi-Heinola, 22.6.2001



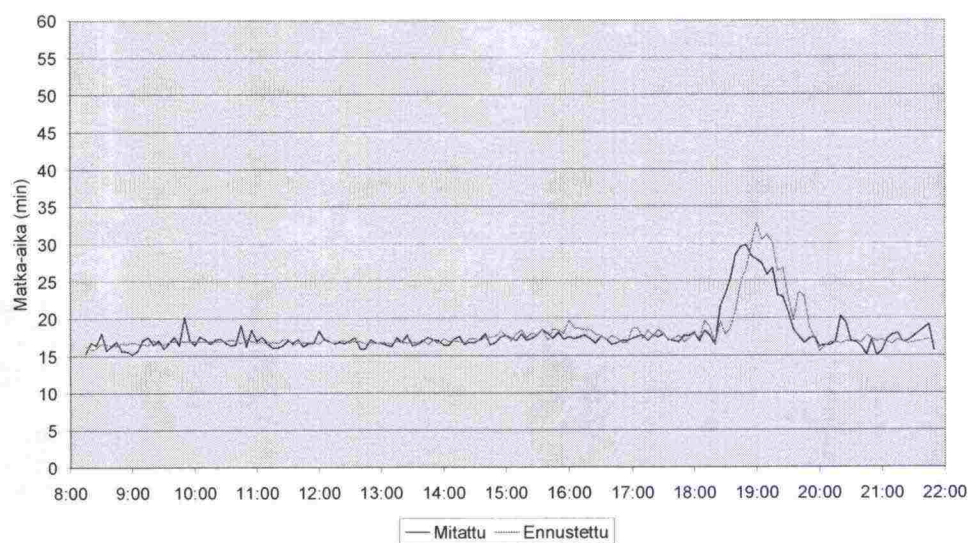
Kymijärvi-Heinola, 22.6.2001



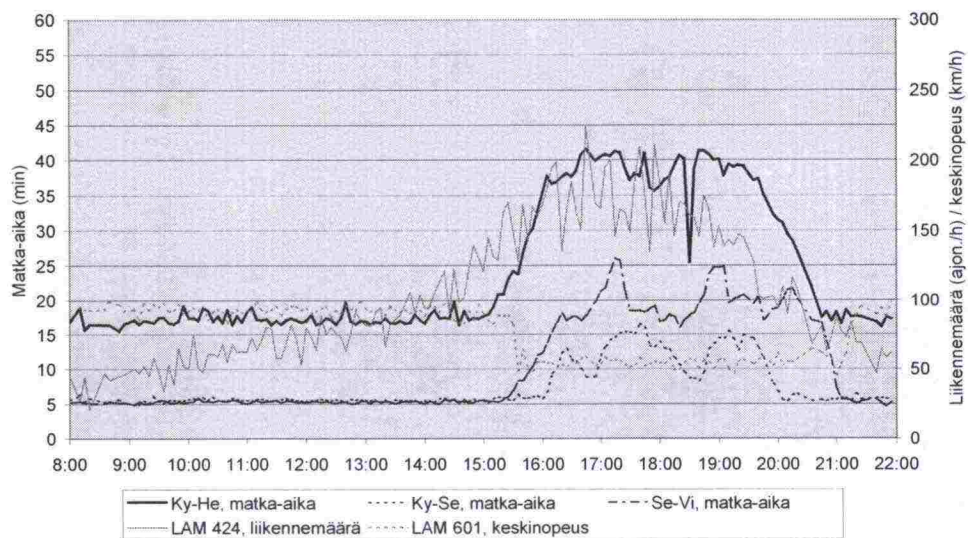
Kymijärvi-Heinola, 29.6.2001



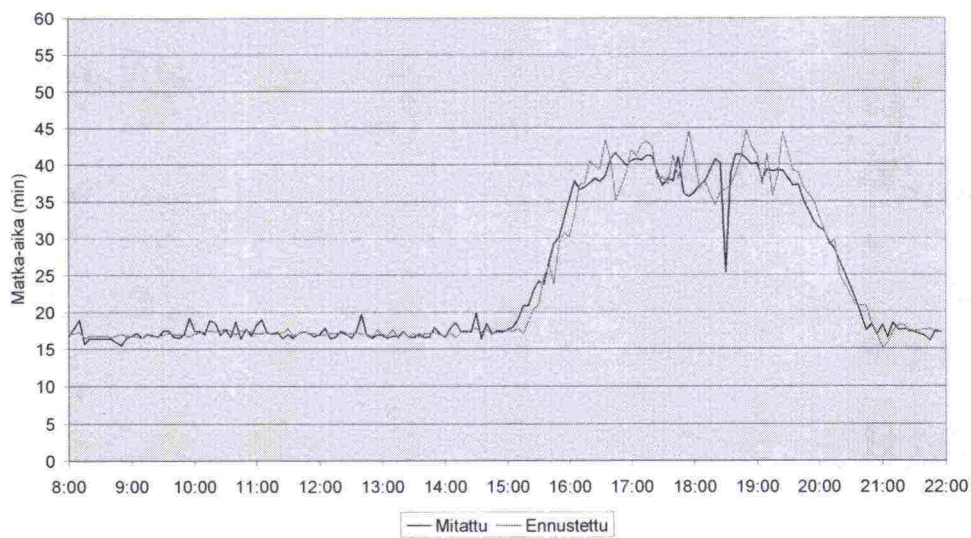
Kymijärvi-Heinola, 29.6.2001



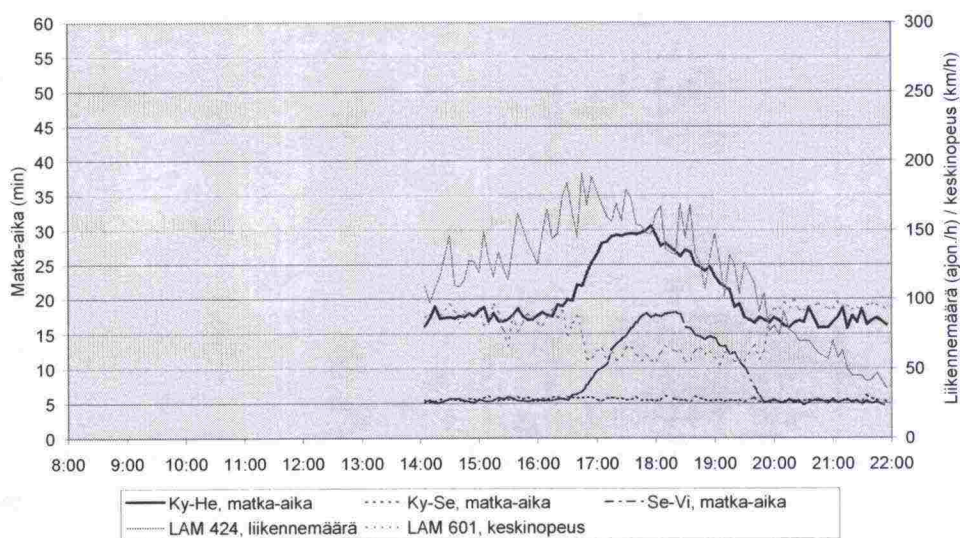
Kymijärvi-Heinola, 6.7.2001



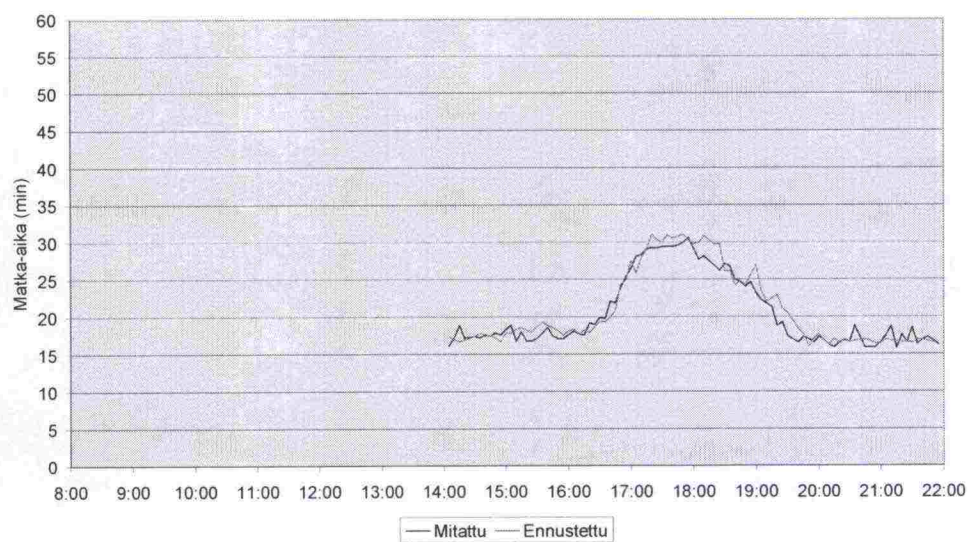
Kymijärvi-Heinola, 6.7.2001



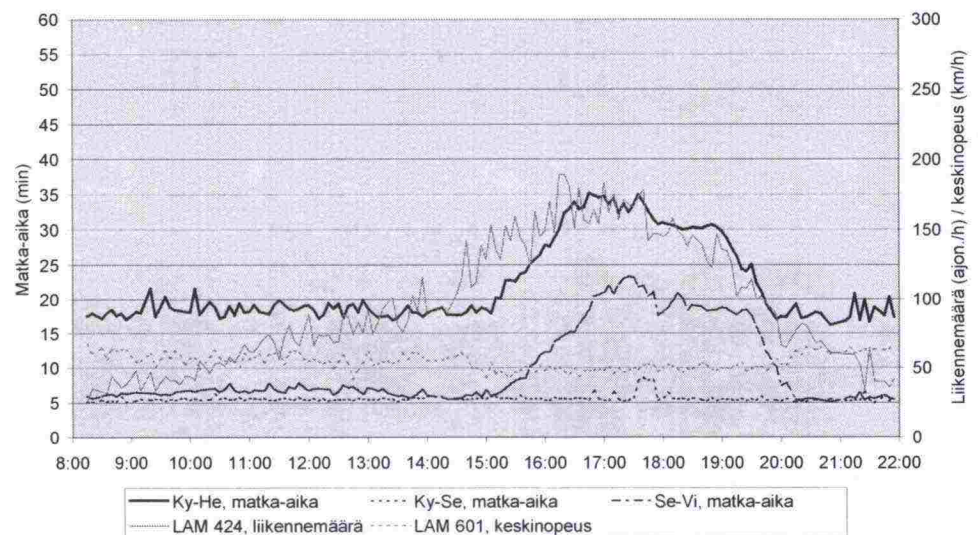
Kymijärvi-Heinola, 20.7.2001



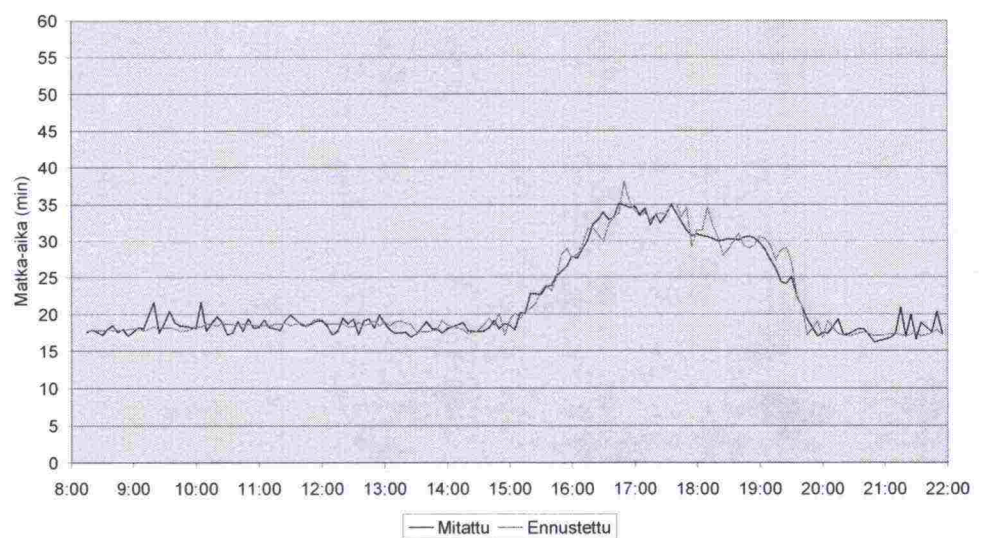
Kymijärvi-Heinola, 20.7.2001



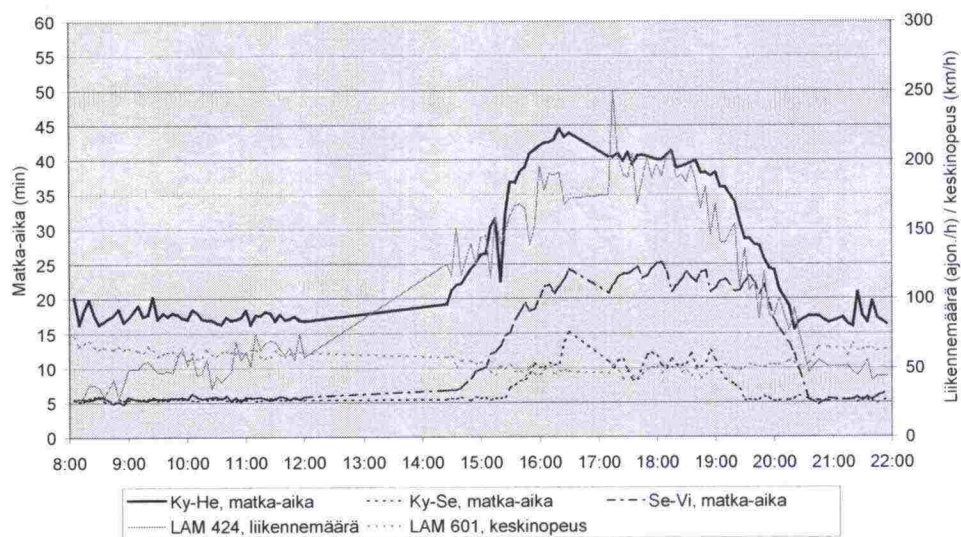
Kymijärvi-Heinola, 27.7.2001



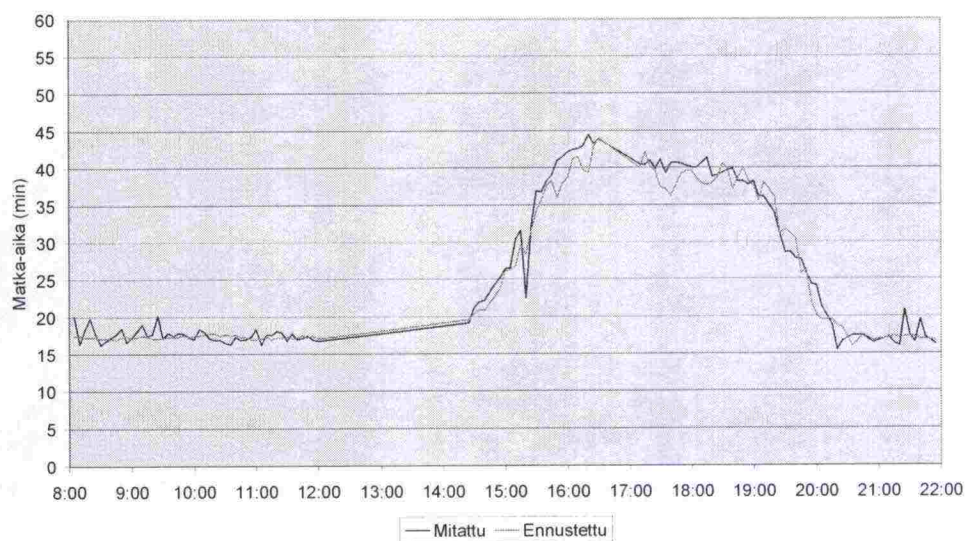
Kymijärvi-Heinola, 27.7.2001



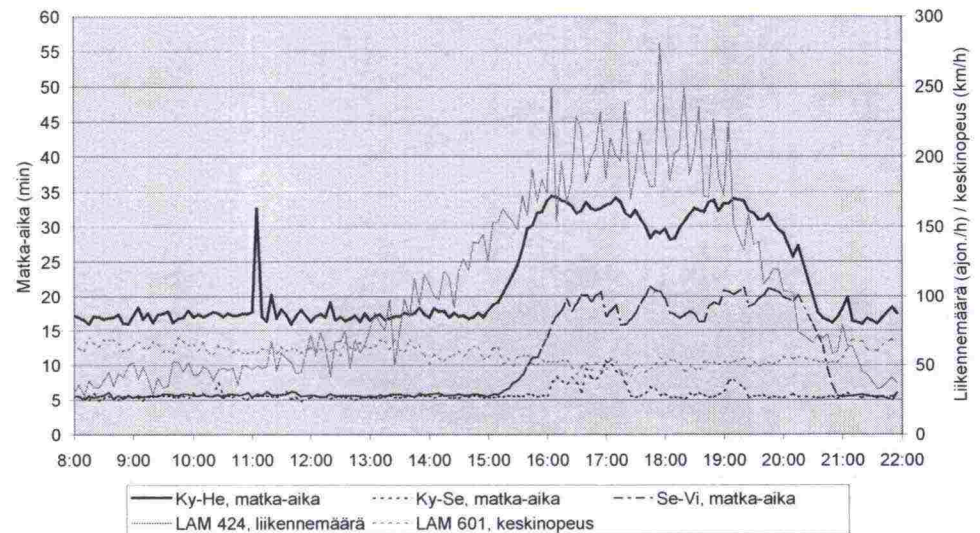
Kymijärvi-Heinola, 3.8.2001



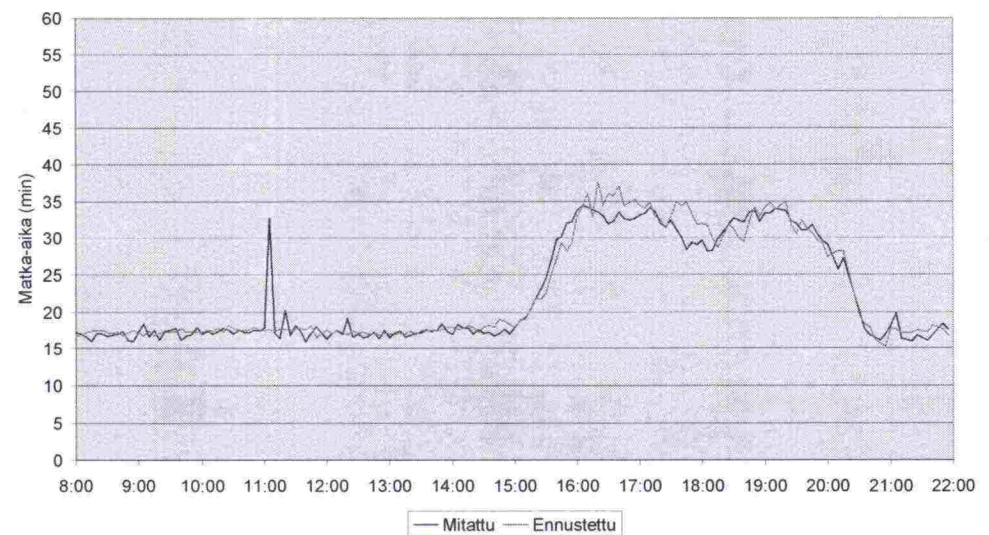
Kymijärvi-Heinola, 3.8.2001



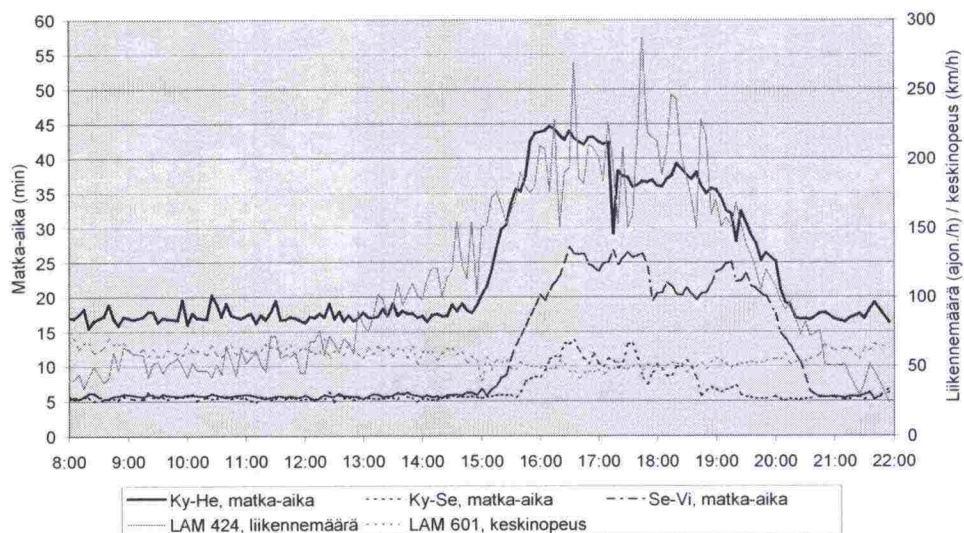
Kymijärvi-Heinola, 17.8.2001



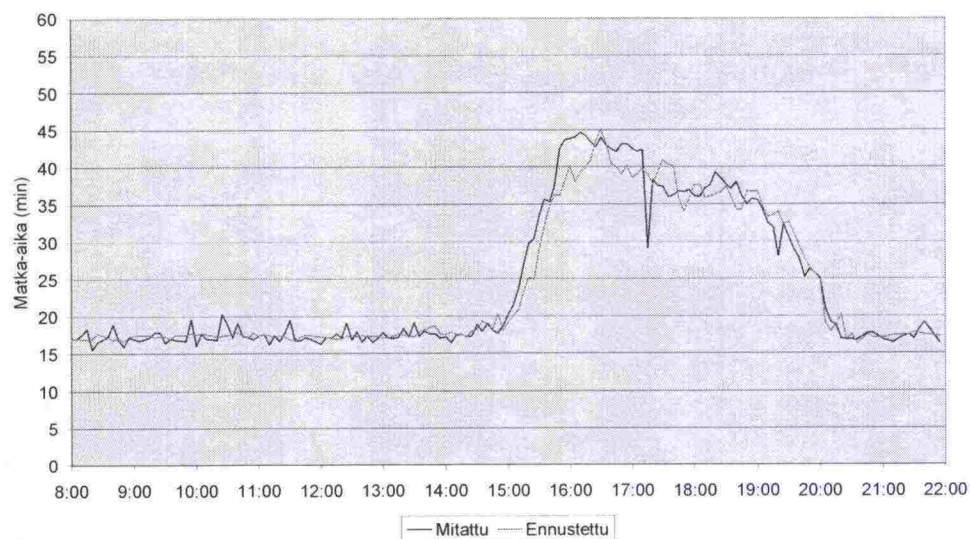
Kymijärvi-Heinola, 17.8.2001



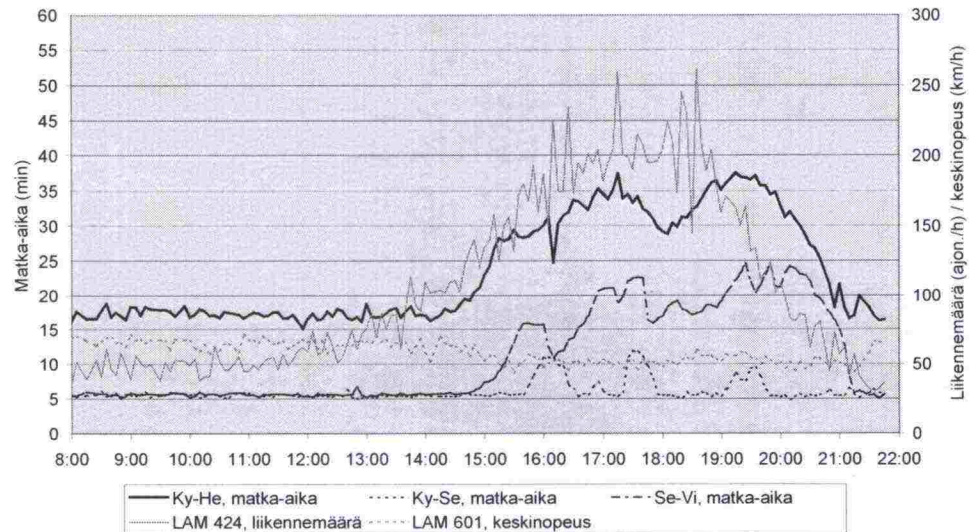
Kymijärvi-Heinola, 24.8.2001



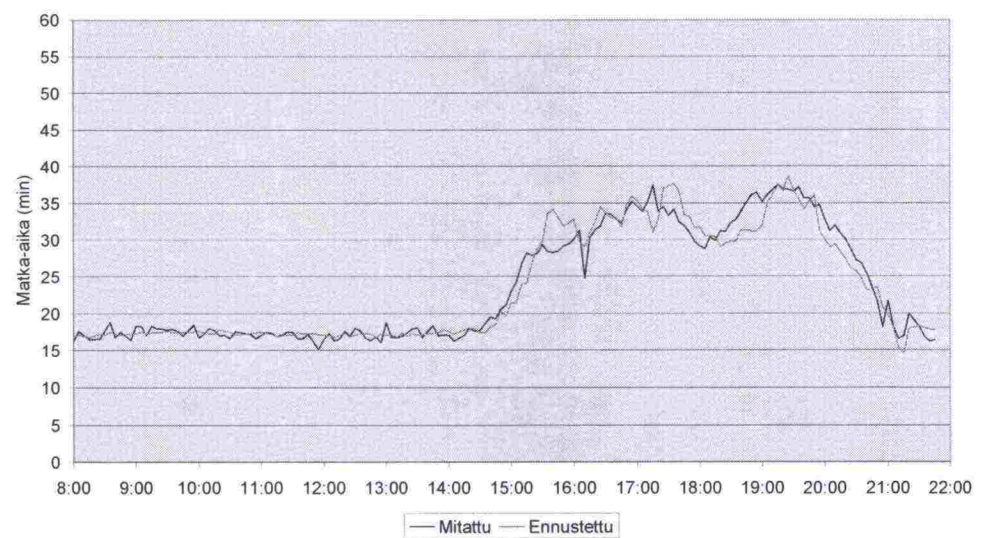
Kymijärvi-Heinola, 24.8.2001



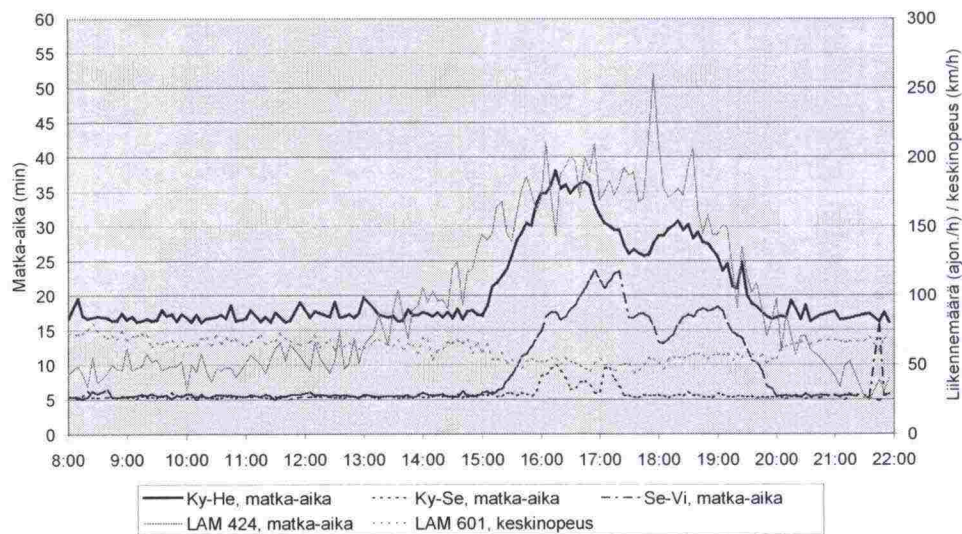
Kymijärvi-Heinola, 31.8.2001



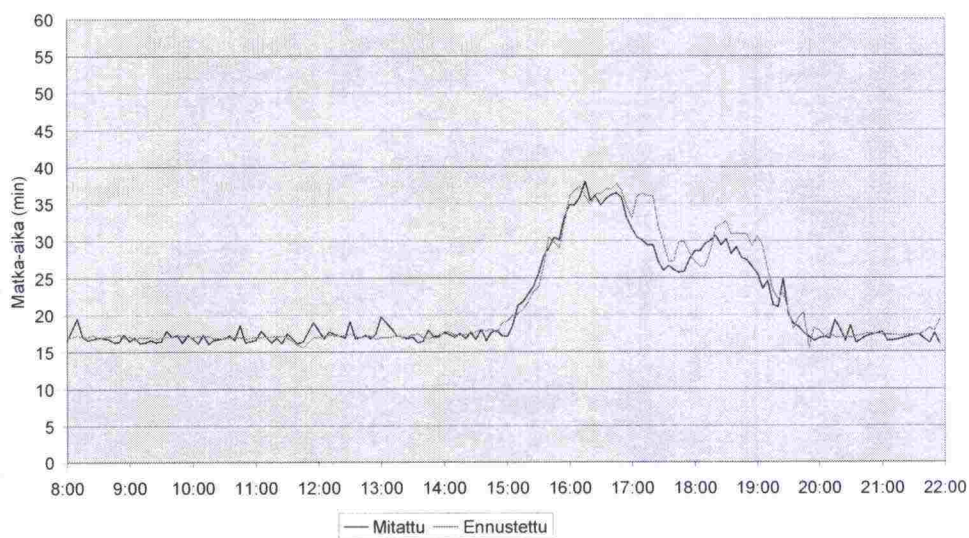
Kymijärvi-Heinola 31.8.2001



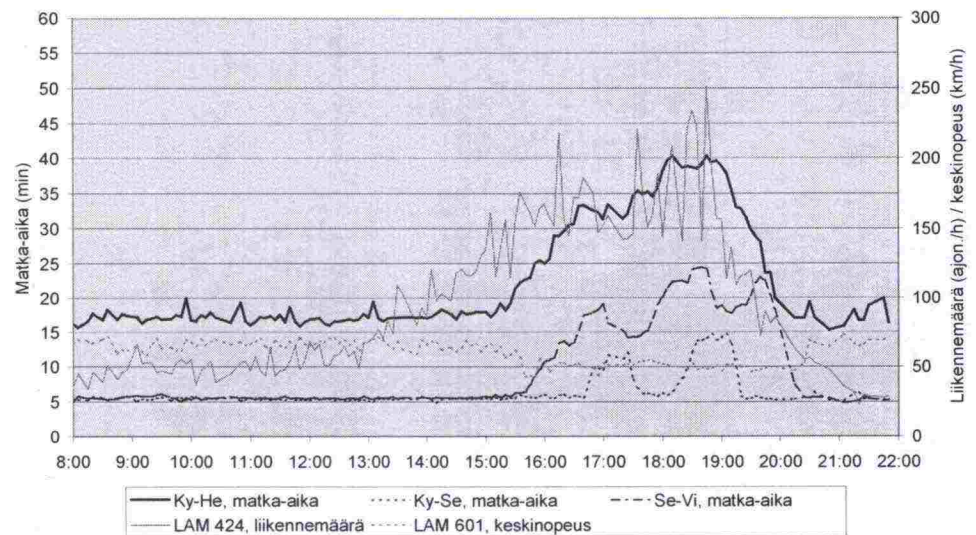
Kymijärvi-Heinola, 7.9.2001



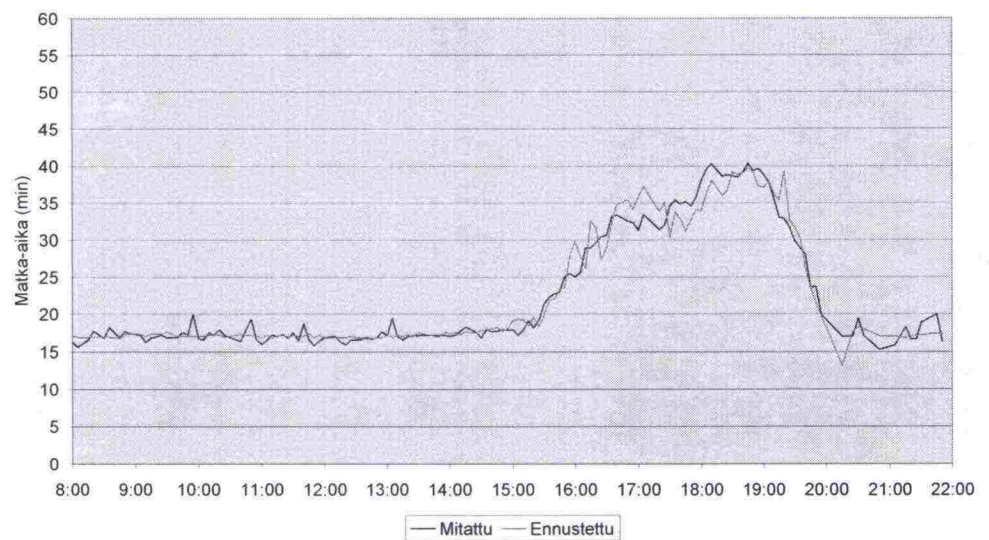
Kymijärvi-Heinola, 7.9.2001



Kymijärvi-Heinola, 14.9.2001

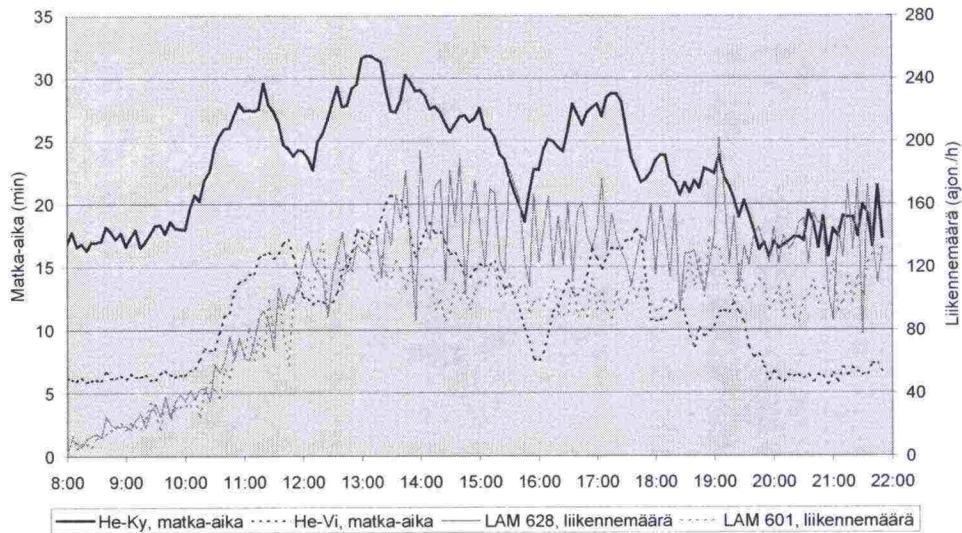


Kymijärvi-Heinola, 14.9.2001

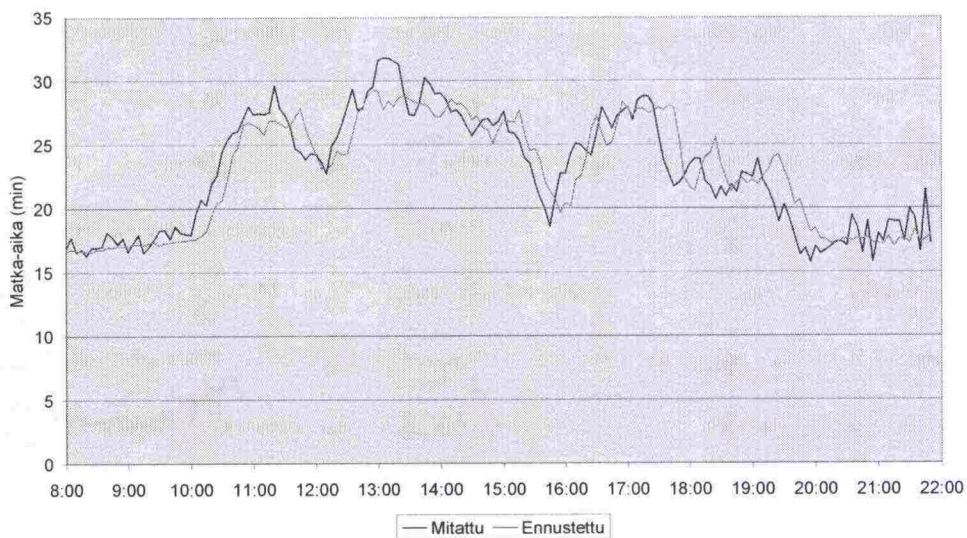


Heinola-Kymijärvi

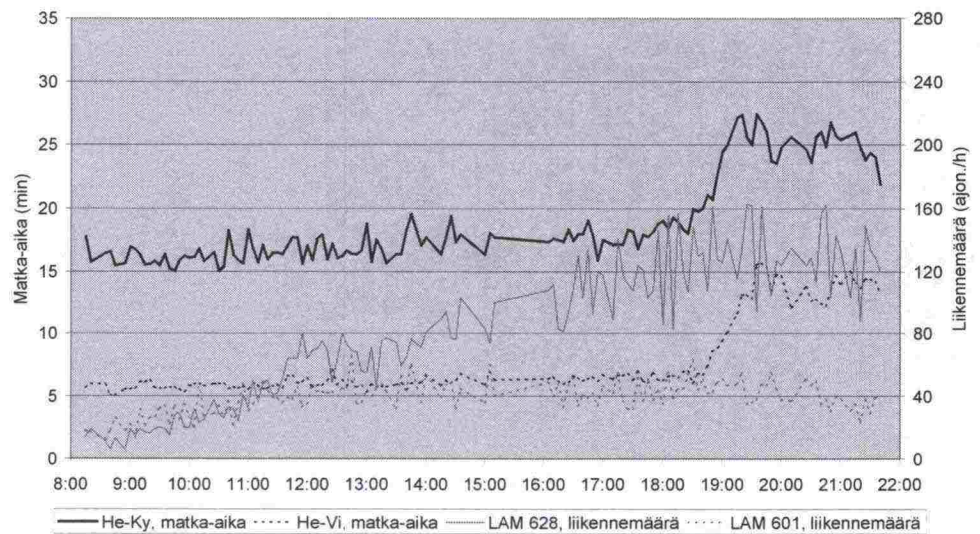
Heinola-Kymijärvi, 25.6.2000



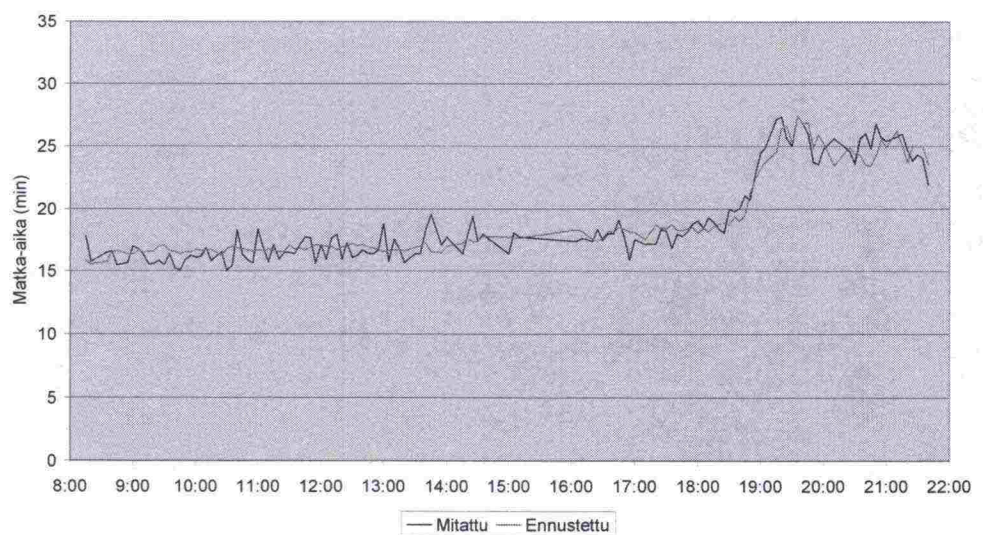
Heinola-Kymijärvi, 25.6.2000



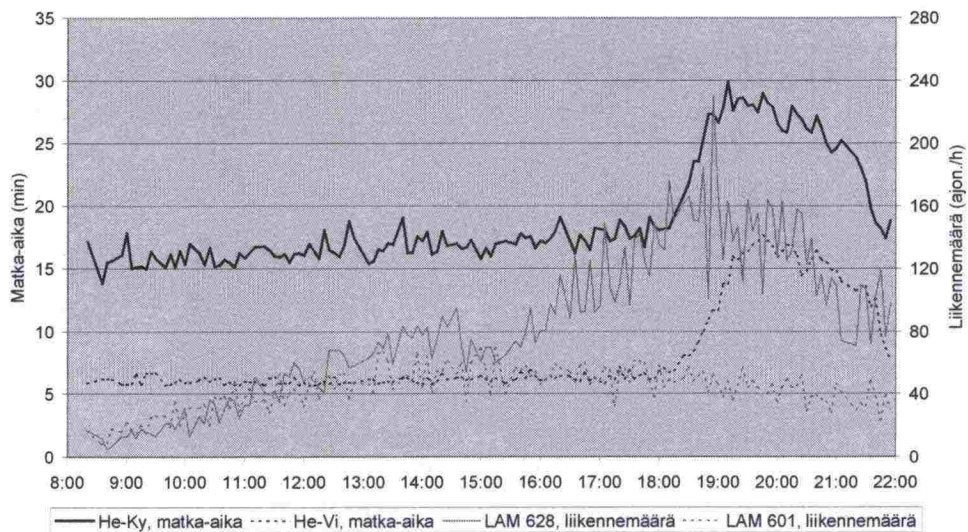
Heinola-Kymijärvi, 8.7.2001



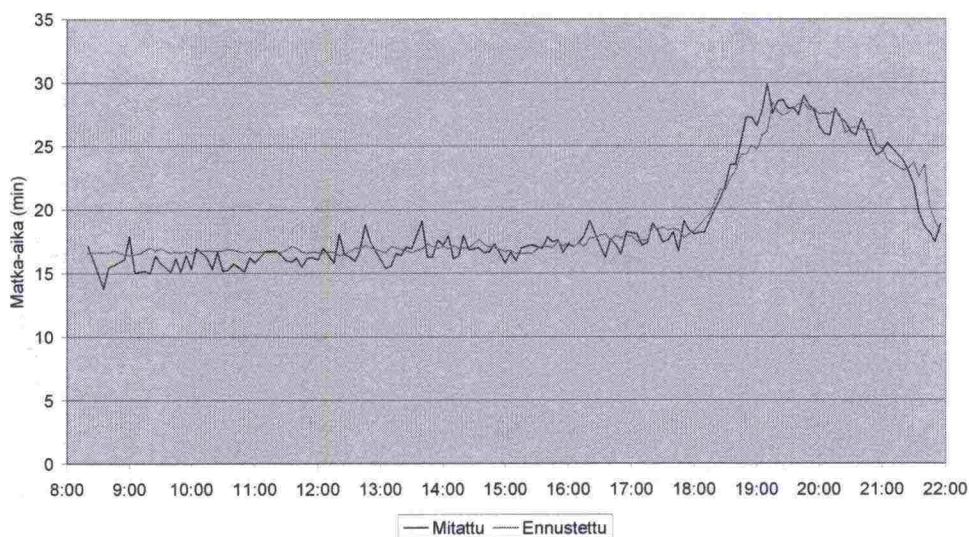
Heinola-Kymijärvi, 8.7.2001



Heinola-Kymijärvi, 15.7.2001



Heinola-Kymijärvi, 15.7.2001



ISSN 1457-9871
ISBN 951-726-897-1
TIEH 3200754